



# Essays on Indices and Matching

## Citation

Shorrer, Ran I. 2015. Essays on Indices and Matching. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467351>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

# Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Essays on Indices and Matching

A dissertation presented

by

Ran Israel Shorrer

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2015

© 2015 Ran Israel Shorrer

All rights reserved.

*Dissertation Advisors:*

**Eric S. Maskin**

**Alvin E. Roth**

*Author:*

**Ran Israel Shorrer**

## **Essays on Indices and Matching**

### **Abstract**

In many decision problems, agents base their actions on a simple *objective index*, a single number that summarizes the available information about objects of choice independently of their particular preferences. The first chapter proposes an axiomatic approach for deriving an index which is *objective* and, nevertheless, can serve as a guide for decision making for decision makers with different preferences. Unique indices are derived for five decision making settings: the Aumann and Serrano (2008) index of riskiness (additive gambles), a novel generalized Sharpe ratio (for a standard portfolio allocation problem), Schreiber's (2013) index of relative riskiness (multiplicative gambles), a novel index of delay embedded in investment cashflows (for a standard capital budgeting problem), and the index of appeal of information transactions (Cabrales *et al.*, 2014). All indices share several attractive properties in addition to satisfying the axioms. The approach may be applicable in other settings in which indices are needed.

The second chapter uses conditions from previous literature on complete orders to generate partial orders in two settings: information acquisition and segregation. In the setting of information acquisition, I show that the partial

order prior independent investment dominance (Cabrales *et al.*, 2013) refines Blackwell’s partial order in the strict sense. In the segregation setting, I show that without the requirement of completeness, all of the axioms suggested in Frankel and Volij (2011) are satisfied simultaneously by a partial order which refines the standard partial order (Lasso de la Vega and Volij, 2014).

In the third and fourth chapters, I turn to examine matching markets. Although no stable matching mechanism can induce truth-telling as a dominant strategy for all participants (Roth, 1982), recent studies have presented conditions under which truthful reporting by all agents is close to optimal (Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Lee, 2011). The third chapter demonstrates that in large, balanced, uniform markets using the Men-Proposing Deferred Acceptance Algorithm, each woman’s best response to truthful behavior by all other agents is to truncate her list substantially. In fact, the optimal degree of truncation for such a woman goes to 100% of her list as the market size grows large. Comparative statics for optimal truncation strategies in general one-to-one markets are also provided: reduction in risk aversion and reduced correlation across preferences each lead agents to truncate more. So while several recent papers focused on the limits of strategic manipulation, the results serve as a reminder that without preconditions ensuring truthful reporting, there exists a potential for significant manipulation even in settings where agents have little information.

Recent findings of Ashlagi *et al.* (2013) demonstrate that in unbalanced random markets, the change in expected payoffs is small when one reverses which side of the market “proposes,” suggesting there is little potential gain from manipulation. Inspired by these findings, the fourth chapter studies the

implications of imbalance on strategic behavior in the incomplete information setting. I show that the “long” side has significantly reduced incentives for manipulation in this setting, but that the same doesn’t always apply to the “short” side. I also show that risk aversion and correlation in preferences affect the extent of optimal manipulation as in the balanced case.

# Contents

Abstract . . . . .	iii
Acknowledgments . . . . .	xiii
<b>1 Consistent Indices</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.1.1 Relation to the Literature . . . . .	8
1.2 Preliminaries . . . . .	11
1.3 Local Aversion to $Q$ -Riskiness . . . . .	16
1.4 The Aversion to AS-Riskiness and the Demand for Gambles .	24
1.5 Global Consistency . . . . .	28
1.6 A Generalized Sharpe Ratio . . . . .	33
1.6.1 Preliminaries . . . . .	36
1.6.2 Results . . . . .	40
1.6.3 The Demand for Market Transactions . . . . .	41
1.7 A Consistent Index of Relative Riskiness . . . . .	43
1.8 Consistent Index of Delay . . . . .	47
1.8.1 Preliminaries . . . . .	48
1.8.2 The Index . . . . .	50
1.8.3 $D$ -Delay Aversion and the Demand for Investment Cash- flows . . . . .	57
1.8.4 Other Properties of $D$ and a Comparison with $Q^{AS}$ . .	59
1.9 A Consistent Index of the Appeal of Information Transactions	62
1.9.1 Preliminaries . . . . .	64
1.9.2 The Index . . . . .	68
1.9.3 The Demand for Information Transactions . . . . .	73

1.9.4	Properties of the Index $A$ . . . . .	74
1.9.5	Prior-Free Implications . . . . .	76
1.10	Discussion . . . . .	80
<b>2</b>	<b>More Complete Incomplete Orders of Informativeness and Segregation</b>	<b>81</b>
2.1	Introduction . . . . .	81
2.2	Preliminaries . . . . .	85
2.3	Results . . . . .	88
2.4	Segregation . . . . .	95
2.4.1	Preliminaries . . . . .	95
2.4.2	Axioms . . . . .	98
2.4.3	Results . . . . .	99
<b>3</b>	<b>Optimal Truncation in Matching Markets</b>	<b>102</b>
3.1	Introduction . . . . .	102
3.2	Matching Markets Background . . . . .	109
3.2.1	Marriage Markets and Stability . . . . .	110
3.2.2	The Men-Proposing Deferred Acceptance Algorithm . . . . .	112
3.2.3	The Preference List Submission Problem for Men . . . . .	113
3.2.4	The Preference List Submission Problem for Women . . . . .	114
3.3	Characterizing Truncation Payoffs . . . . .	119
3.3.1	Truncation in Two Stages: Match, then Divorce . . . . .	120
3.3.2	Truncation under General Beliefs . . . . .	121
3.3.3	Truncation under $\mathcal{M}$ -Symmetric Beliefs . . . . .	123
3.3.4	Optimal Truncation in Large Markets . . . . .	125
3.4	Truncation Equilibria and Welfare . . . . .	129
3.4.1	Equilibria under General Preferences . . . . .	130
3.4.2	Equilibria in Uniform Markets . . . . .	132
3.4.3	Simulations: Finding a Symmetric Equilibrium . . . . .	136
3.5	Truncation and Risk Aversion . . . . .	140
3.6	Correlated Preferences . . . . .	145
3.6.1	Perfectly Correlated Preferences . . . . .	146
3.6.2	Partially Correlated Preferences . . . . .	146



3.6.3	Noisy Preferences . . . . .	148
3.7	Discussion and Conclusion . . . . .	151
<b>4</b>	<b>Strategic Behavior in Unbalanced Matching Markets</b>	<b>153</b>
4.1	Introduction . . . . .	153
4.2	Preliminaries . . . . .	156
4.2.1	Marriage Markets and Stability . . . . .	156
4.2.2	The Men-Proposing Deferred Acceptance Algorithm . . . . .	158
4.2.3	The Preference List Submission Problem . . . . .	160
4.3	Optimal Truncation in Unbalanced Markets . . . . .	164
4.3.1	The Case of More Women Than Men . . . . .	166
4.3.2	The Case of Fewer Women Than Men . . . . .	173
4.4	Other Aspects Impacting the Optimal Level of Truncation . . . . .	177
4.4.1	Truncation and Risk Aversion . . . . .	177
4.4.2	Truncation and Correlated Preferences . . . . .	180
4.5	Conclusion . . . . .	182
	<b>References</b>	<b>185</b>
<b>A</b>	<b>Appendix to Chapter 1</b>	<b>195</b>
A.1	Proofs . . . . .	195
A.1.1	Fact A.1 and Lemmata A.1 and A.2 . . . . .	195
A.1.2	Theorem 1.2 . . . . .	197
A.1.3	Theorem 1.3 . . . . .	198
A.1.4	Proposition 1.2 . . . . .	200
A.1.5	Theorems 1.4 and 1.5 . . . . .	200
A.1.6	Theorem 1.7 . . . . .	201
A.1.7	Theorem 1.8 . . . . .	202
A.1.8	Lemma 1.4 . . . . .	205
A.1.9	Theorem 1.10 . . . . .	206
A.1.10	Theorem 1.11 . . . . .	209
A.1.11	Theorem 1.14 . . . . .	210
A.1.12	Propositions 1.10 and 1.11 . . . . .	213
A.1.13	Theorem 1.15 . . . . .	213

A.1.14 Theorem 1.16 . . . . .	215
A.1.15 Theorem 1.17 . . . . .	217
A.1.16 Theorem 1.20 . . . . .	218
A.1.17 Theorem 1.21 . . . . .	219
A.1.18 Theorem 1.22 . . . . .	220
A.1.19 Theorem 1.24 . . . . .	221
<b>B Appendix to Chapter 3</b>	<b>223</b>
B.1 Proofs . . . . .	223

## List of Figures

- 1.1 The figure depicts the two dimensional space of  $2 \times 2$  information structures. These matrices could be written as  $\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$ , where both  $x$  and  $y$  are in  $[0, 1]$ . In the figure,  $x$  is represented by the horizontal axis and  $y$  is represented by the vertical axis. The shaded area are the matrices which represent information structures which are dominated by  $\alpha_1$  in the Blackwell sense. The point  $\alpha_2$  is outside the shaded area. . . . . 79
- 2.1 The figure depicts the two dimensional space of  $2 \times 2$  information structures. These matrices could be written as  $\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$ , where both  $x$  and  $y$  are in  $[0, 1]$ . In the figure,  $x$  is represented by the horizontal axis and  $y$  is represented by the vertical axis. The shaded area are the matrices which represent information structures which are dominated by  $\alpha_1$  in the Blackwell sense. The point  $\alpha_2$  is outside the shaded area. . . . . 91
- 3.1 **Simulation Results for Truncation Payoffs.** The graphs display  $(N+1)$ – an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, N\}$  and submitting these preferences to *MP-DA*. Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. Markets are of size 10, 100, 1,000 and 10,000. . . . . 128

3.2	Woman $w$ 's expected payoff in a uniform market as a function of her truncation point when women $\mathcal{W} \setminus \{w\}$ truncate at $j$ . $N = 30$ . Iterations = 1,000,000. . . . .	136
3.3	Optimal truncation point of woman $w$ in a uniform market as a function of the common truncation point $j$ of women $\mathcal{W} \setminus \{w\}$ . $N = 30$ . Iterations = 10,000,000. . . . .	138
3.4	$k$ -Truncation is equivalent to $(k + 1)$ -truncation followed by an extra gamble: divorcing man $k + 1$ . . . . .	142
3.5	The graphs display $(N + 1)$ – an individual woman's expected partner rank from truncating her list at each point $k \in \{0, \dots, N\}$ and submitting these preferences to <i>MP-DA</i> . Preference lists of men are uniformly random, and lists for women are randomly generated using the procedure described in the text. Payoffs are averaged over 100,000 draws. . . . .	150
4.1	<b>Simulation Results for Truncation Payoffs.</b> The graphs display $(N + 1)$ – an individual woman's expected partner rank from truncating her list at each point $k \in \{0, \dots, N\}$ and submitting these preferences to <i>MP-DA</i> . Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. Markets are of size 10, 100, 1,000 and 10,000. . . . .	166
4.2	<b>Simulation Results for Truncation Payoffs.</b> The graphs display 101– an individual woman's expected partner rank from truncating her list at each point $k \in \{0, \dots, 100\}$ and submitting these preferences to <i>MP-DA</i> . Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. All markets have 100 men, and the number of women varies between 90, 95, 99, 101, 105 and 110. . . . .	168

4.3	<b>Simulation Results for Truncation Payoffs.</b>	In markets with 1000 men and 999 or 1001 women, the graph displays 1001–an individual woman’s expected partner rank from truncating her list at each point $k \in \{0, \dots, 1000\}$ and submitting these preferences to <i>MP-DA</i> . Preference lists of the other agents are selected, uniformly at random, from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws.	177
-----	---	--	-----

## Acknowledgments

Thanks to Mitra Akhtari, Itai Ashlagi, Nageeb Ali, Robert Aumann, Dany Bahar, Peter Coles, Oren Danieli, William Diamond, Dean Foster, Mira Frick, Drew Fudenberg, Ben Golub, Jerry Green, Sergiu Hart, Avinatan Hassidim, Johannes Hörner, Guido Imbens, Nicole Immorlica, Yash Kanoria, Divya Kirti, Peter Klibanoff, Scott Kominers, David Laibson, Jacob Leshno, Annie Liang, Fabio Maccheroni, Mark Machina, Eric Maskin, Moti Michaeli, Jenő Pál, Ron Peretz, Assaf Romm, Al Roth, Ben Roth, Amnon Schreiber, Philipp Strack, Tomasz Strzalecki, Neil Thakral, William Thomson and Yufei Zhao. Financial support from the Geneva Association is gratefully acknowledged.

To Daniel

# Chapter 1

## Consistent Indices

### 1.1 Introduction

In many decision problems, agents base their actions on a simple *objective index*, a single number that summarizes the available information about objects of choice and does not depend on the agent’s particular preferences.<sup>1</sup> Agents might choose to do this due to difficulties in attaining and interpreting information, or due to an overabundance of useful information. For example, the *Sharpe ratio* (Sharpe, 1966), the ratio between the expected net return and its standard deviation, is frequently used as a performance measure for portfolios (Welch, 2008; Kadan and Liu, 2014).

This paper proposes an axiomatic approach for deriving an index that is *objective* and, nevertheless, can serve as a guide for decision making for decision makers with different preferences. The approach is unifying and may be used in a variety of decision making settings. I present five applications:

---

<sup>1</sup>As shown by Luca (2011), for the case of online restaurant star ratings.



for a setting of *additive gambles*, which like lottery tickets change the baseline wealth of the owner independently of its level (an *index of riskiness*); for a standard *portfolio allocation* problem (a *generalized Sharpe ratio*); for a setting of *multiplicative gambles*, which change the wealth of the owner proportionally to its baseline level (an *index of relative riskiness*); for a standard *capital budgeting* problem (an index of the *delay* embedded in investment cashflows); and for a setting of *information acquisition* by investors in an Arrowian (Arrow, 1972) environment (an *index of appeal of information transactions*). In each of the settings I study, a unique index emerges that is theoretically appealing and often improves upon commonly used indices. The approach may be applicable in other settings in which indices are needed.<sup>2</sup>

In my setting, agents choose whether to accept or reject a transaction (a gamble, a cashflow, etc.). The starting point of this paper is a given decision problem and the requirement that (at least) small decisions can be made based on the index. This is the content of the *local consistency* axiom. The axiom states, roughly, that all agents can make acceptance and rejection decisions for small, “local,” transactions using the index and a *cut-off* value (which is the only parameter that depends on their preferences), without knowing other details about the transaction, so that the outcomes of their decisions will mirror the outcomes they would achieve by optimizing when possessing detailed knowledge about the transaction.

Even though transactions are complex and multidimensional, I show that a numeric, single dimensional, index can summarize all the decision-relevant

---

<sup>2</sup>A particular setting which seems promising in this regard is the measurement of inequality, which has many similarities to the setting of risk (Atkinson, 1970).

information for small transactions. I thus view local consistency as a minimal requirement for an index to be a useful guide for decision making, and, as I show, it is indeed satisfied by many well-known indices in various decision making problems. However, while this property is desirable, I show that many indices that have it also have normatively undesirable properties.<sup>3</sup> The Sharpe ratio, for example, has such property outside the domain of normal distributions. As shown in Example 1.10, the Sharpe ratio is not, in general, monotonic with respect to first order stochastic dominance outside that domain (Hodges, 1998).<sup>4</sup>

A second criterion for assessing the validity of an index, *global consistency*, is therefore suggested. Global consistency extends local consistency by making restrictions over large transactions, but it is actually quite a weak restriction. Nevertheless, the combination of local and global consistency turns out to be powerful. In the various decision making problems which are discussed below, it pins down a unique order over transactions that has several desirable properties in addition to local and global consistency.<sup>5</sup> Since I use results from the setting of additive gambles in my treatment of other decision making

---

<sup>3</sup>As stated here, the result follows trivially given the existence of one locally consistent index, as one could change the values of large transactions without changing those of small, local, ones. The exact statement makes further technical requirements which disqualify such indices.

<sup>4</sup>This undesirable property is related to the fact that this index depends only on the first two moments of the distribution. These moments are sufficient statistic for a normal distribution, and therefore basing an index on them solely may be reasonable if returns are assumed (or known) to be normally distributed. This assumption, however, is often rejected in empirical tests in settings where the Sharpe ratio is used in practice (e.g. Fama, 1965; Agarwal and Naik, 2000; Kat and Brooks, 2001). Moreover, a large body of literature documents the importance of higher order moments for investment decisions (e.g. Kraus and Litzenberger, 1976; Kane, 1982; Harvey and Siddique, 2000; Barro, 2006, 2007; Gabaix, 2008).

<sup>5</sup>To be precise, additional mild conditions are required as well.

environments, I begin by reviewing this setting and cover it in detail in order to illustrate the general concepts.

The approach I take is different from the standard decision theoretic approach. I start with a given *objective index* – a function that assigns to each transaction some number, independently of any agent-specific characteristics. In the case of additive gambles, a higher number is associated with a higher level of riskiness. As different functions induce different orders, for a given index  $Q$ , I refer to the  $Q$ -riskiness of a gamble. Only then I define the aversion to  $Q$ -riskiness. I define the relation *locally at least as averse to  $Q$ -riskiness* as follows: agent  $u$  with wealth  $w$  is locally at least as averse to  $Q$ -riskiness as agent  $v$  with wealth  $w'$  if, for all gambles with small support (defined precisely in Section 1.3),<sup>6</sup> when  $u$  at  $w$  accepts any small gamble with a certain level of  $Q$ -riskiness,  $v$  at  $w'$  accepts all small gambles which are significantly less  $Q$ -risky. This definition assumes a certain kind of consistency between the index and the aversion to the property it evaluates, as it implies that agents that are less  $Q$ -riskiness averse would accept  $Q$ -riskier gambles. This approach is the dual of the standard approach, since instead of starting with an ordering over preferences and asserting that risk is “*what risk-aversers hate*” (Machina and Rothschild, 2008), I start with an ordering over the objects of choice (an index of riskiness  $Q$ ) and derive from it judgments on preferences ( $Q$ -riskiness aversion).

---

<sup>6</sup>The need to restrict attention to small supports is nicely illustrated by a discussion Samuelson (1963) describes having with Stanislaw Ulam. Samuelson (1963) quotes Ulam as saying “*I define a coward as someone who will not bet when you offer him two-to-one odds and let him choose his side,*” to which he replied “*You mean will not make a sufficiently small bet (so that the change in the marginal utility of money will not contaminate his choice).*”

In Section 1.3, I show that if  $Q$  is a locally consistent index which satisfies an additional mild condition, then the relation “at least as averse to  $Q$ -riskiness” induces the same order as the classic *coefficient of absolute risk aversion* (ARA, Pratt, 1964; Arrow, 1965, 1971). This property is shown to be satisfied by several well-known indices. However, it is also satisfied by many other indices, including ones that are not monotonic with respect to first order stochastic dominance (Hanoch and Levy, 1969; Hadar and Russell, 1969; Rothschild and Stiglitz, 1970).

As local consistency is insufficient for pinning down normatively acceptable indices, in Section 1.4 I propose a second criterion — the *generalized Samuelson property*. An index of riskiness has this property when no agent accepts a large gamble of a certain degree of riskiness if he rejects small ones of the same degree of riskiness at any wealth level, and no agent rejects a large gamble of a certain degree of riskiness if he accepts small ones of the same degree of riskiness at any wealth level. I also show that no agent whose risk tolerance (the inverse of the coefficient of absolute risk aversion) is always higher than the AS riskiness of  $g$  will reject  $g$ , and no agent whose risk tolerance is always lower will accept it. Given an empirical range of the degrees of risk aversion in a population, the model provides advice to individuals and policy makers based on the index. It also allows researchers a simple way to estimate bounds on the degree of risk aversion in the population from observations of acceptance and rejection of different gambles.

In Section 1.5, I show that the generalized Samuelson property can be replaced by a weaker condition that involves pairs of agents — *global consistency*. I say that one agent is *globally at least as averse to  $Q$ -riskiness*

as another agent, if he is locally at least as averse to  $Q$ -riskiness at any two arbitrary wealth levels. In the additive gambles setting, global consistency requires that if two agents can be compared using this partial order, then the more  $Q$ -riskiness averse agent rejects gambles which are riskier than ones rejected by the other agent. Note that the partial order on preferences which is used to make this requirement of consistency is defined using the index  $Q$ , and not based on preexisting notions of risk aversion. Global consistency is a weak requirement, in the sense that it imposes no restriction for the (common) case of a pair of agents who cannot be compared using this partial order. However, I show that with additional mild conditions, the Aumann and Serrano (2008) index of riskiness, which is monotonic with respect to stochastic dominance, is the unique index that satisfies local consistency and global consistency.

Section 1.6 addresses the ranking of performance of a market portfolio in the presence of a risk-free asset. One well known index of performance is the Sharpe ratio (Sharpe, 1966), the ratio between the expected net return and its standard deviation. Using the approach from Section 1.5 I derive the *generalized Sharpe ratio*, where the role of the standard deviation is taken by the Aumann-Serrano (AS) index. This index of performance coincides with the Sharpe ratio on the domain of normal distributions but differs from it in general.<sup>7</sup> Unlike the Sharpe ratio, it is monotonic with respect to stochastic dominance, even when the risky return is not normally distributed, and it satisfies other desirable properties.

Section 1.7 covers the setting of multiplicative gambles. The results are

---

<sup>7</sup>The index is increasing in odd distribution moments and decreasing in even ones.

quite analogous to those of the additive gambles setting. The role of ARA is replaced by the *coefficient of relative risk aversion* (RRA). I show that with mild conditions, the index of relative riskiness of Schreiber (2013) is the unique index which satisfies local consistency and global consistency (or the generalized Samuelson property).

Section 1.8 considers a capital budgeting setting. Agents are proposed *investment cashflows*, opportunities of investment for several periods with return at later times. I label indices for this setting as indices of *delay*. Paralleling results in previous sections, I show that local consistency, combined with additional mild conditions, ensures that the local aversion to delay, as defined by an index, is ordinally equivalent to the instantaneous discount rate. Adding the requirement of global consistency (or the generalized Samuelson property) is then shown to pin down a novel index for the delay embedded in investment cashflows. The index is continuous and monotonic with respect to *time dominance* (Bøhren and Hansen, 1980; Ekern, 1981), a partial order on cashflows in the spirit of stochastic dominance.

Section 1.9 treats the setting of information acquisition by investors facing a standard investment problem (Arrow, 1972). I show that the local taste for informativeness, as defined by the index, coincides with the inverse of ARA for any index which satisfies local consistency and another mild condition. These include Cabrales *et al.* (2013) and Cabrales *et al.* (2014), but also indices which have a normatively undesirable property: they are not monotonic with respect to Blackwell's (1953) partial order.<sup>8</sup> I then show that the index

---

<sup>8</sup>One information structure dominates the other in the sense of Blackwell if it is preferred to the other by all decision makers for all decision making problems.

of Cabrales *et al.* (2014) is the unique index which satisfies the additional requirement of global consistency.

### 1.1.1 Relation to the Literature

Apart from serving as input in decision making processes, indices are also used to limit the discretion of agents by regulators (Artzner, 1999) or when decision rights are being delegated (Turvey, 1963). For example, a mutual fund manager may be required to invest in bonds that are rated AAA. Similarly, credit decisions are frequently based on a credit rating, a number that is supposed to summarize relevant financial information about an individual. Indices are also used in empirical studies in order to evaluate complex, multidimensional, attributes. Examples include the cost of living (Diewert, 1998), segregation (Echenique and Fryer Jr., 2007), academic influence (Palacios-Huerta and Volij, 2004; Perry and Reny, 2013), market concentration (Herfindahl, 1950), the upstreamness of production and trade flows (Antràs *et al.*, 2012), contract intensity in production (Nunn, 2007), centrality in a network (Bonacich, 1987), inequality (Yitzhaki, 1983; Atkinson, 1983), poverty (Atkinson, 1987), risk and performance (Sharpe, 1966; Artzner *et al.*, 1999), political influence (Shapley and Shubik, 1954; Banzhaf III, 1964), and corruption perceptions (Lambsdorff, 2007).

Although indices are used extensively in economic research and in practice, in many cases the index is not carefully derived from theory. Even in cases where they make theoretical sense in a specific setting, they are often used in larger domains. For example, risk has been evaluated using numerous indices

including the standard deviation of returns, the Sharpe ratio, value at risk (VaR), variance over expected return and the coherent measures of Artzner *et al.* (1999).<sup>9</sup> Some of these indices, like the Sharpe ratio, suffer from a severe normative drawback: they are not monotonic with respect to first order stochastic dominance outside specific domains.<sup>10</sup> That is, increasing a gamble's value in every state of the world does not necessarily lead the index to deem it less risky. Different indices have other undesirable properties. For example, some indices are not continuous, which makes them hard to estimate empirically. Some indices, like VaR, are independent of outcomes in the tails. Finally, and key to this paper, some of the indices are not locally consistent,<sup>11</sup> so they may not be used to guide decisions. My approach is to consider fairly general settings and concentrate on consistency.

This paper contributes to the growing literature, pioneered by Aumann and Serrano (2008), which identifies objective indices for specific decision making problems. For additive gambles, Aumann and Serrano present an objective index of riskiness, based on a small set of axioms, including centrally a “duality axiom,” which requires a certain kind of consistency. Roughly speaking, it asserts that (uniformly) less risk-averse individuals accept riskier gambles.<sup>12</sup> Importantly, their definition of risk aversion takes the traditional view, and does not refer to risk as defined by the index. Foster and Hart

---

<sup>9</sup>Even though all of the above indices are meant to measure “*risk*,” they were derived with different decision making problems in mind: some take the point of view of a regulator, and others of an investor; some assume the existence of a risk-free asset and others do not; some allow agents to adjust their level of investment, and others assume indivisible assets.

<sup>10</sup>See Example 1.10.

<sup>11</sup>See Example 1.1.

<sup>12</sup>Agent  $i$  uniformly no less risk-averse than agent  $j$  if whenever  $i$  accepts a gamble at some wealth,  $j$  accepts that gamble at any wealth.



(FH, 2009) present a different index of riskiness with an operational interpretation.<sup>13</sup> Their index identifies for every gamble the critical wealth level below which it becomes “risky” to accept the gamble.<sup>14</sup> Schreiber (2013) uses insights from this literature to develop an *index of relative riskiness* for multiplicative gambles. Cabrales *et al.* (2013) and Cabrales *et al.* (2014) treat the setting of information acquisition and the *appeal* of different *information transactions* for investors.

My approach provides a unifying framework for the decision making problems mentioned above, and it can also be applied to new settings. It provides the first axiomatization for the index of delay and for the generalized Sharpe ratio. All of the indices share several desirable properties, such as monotonicity (e.g., with respect to stochastic dominance) and continuity. The generalized Sharpe ratio, one of the two novel indices presented here, is monotonic with respect to *stochastic dominance in the presence of a risk-free asset* (Levy and Kroll, 1978), the analogue of stochastic dominance, of the first and second degree. The index of delay is monotonic with respect to time dominance (Bøhren and Hansen, 1980; Ekern, 1981), the analogous partial order on cashflows.

The index of delay is closely related to a well-known measure of delay which is used in practice: the *internal rate of return* (IRR). I discuss this relation as well as the close connection of the index to the AS index of riskiness. Like the generalized Sharpe ratio, this index treats a decision making environment

---

<sup>13</sup>Homm and Pigorsch (2012b) provide an operational interpretation of the Aumann–Serrano index of riskiness.

<sup>14</sup>Hart (2011) later demonstrated that both indices also arise from a comparison of acceptance and rejection of gambles.

which has not yet been treated by the recent literature on indices for decision problems. These applications therefore underscore a strength of the proposed approach: indices emerge from the same requirements in different decision making environments.

This paper also contributes to the literature that attempts to extend the partial order of Blackwell by restricting the class of decision problems and agents under consideration (e.g. Persico, 2000; Athey and Levin, 2001; Jewitt, 2007). Both Cabrales *et al.* (2014) and Cabrales *et al.* (2013) treat an investment decision making environment with a known, common and fixed prior. The order induced by their indices depends on this prior; there exists pairs of information transactions which are ranked differently depending on the prior selected. But an analyst cannot always observe the relevant prior. Subsection 1.9.5 asks whether the index I derive has prior-free implications for the way information transactions are ranked, which go beyond monotonicity in Blackwell's order and in price. The answer is shown to be positive: there exist pairs of information structures such that neither dominates the other in the sense of Blackwell, and when priced identically, one is ranked higher than the other by the index of appeal of information transactions for any prior distribution. A similar result is shown by Shorrer (2015) for the index of Cabrales *et al.* (2013).

## 1.2 Preliminaries

In this section I provide some notation which will be required for the next sections.

A *gamble*  $g$  is a real-valued random variable with positive expectation and some negative values (i.e.,  $\mathbb{E}[g] > 0$  and  $\Pr\{g < 0\} > 0$ ); for simplicity, I assume that  $g$  takes finitely many values.  $\mathcal{G}$  is the collection of all such gambles. For any gamble  $g \in \mathcal{G}$ ,  $L(g)$  and  $M(g)$  are respectively the maximal loss and gain from the gamble that occur with positive probability. Formally,  $L(g) := \max \text{supp}(-g)$  and  $M(g) := \max \text{supp}(g)$ .

$\mathcal{G}_\epsilon$  is the class of gambles with support contained in an  $\epsilon$ -ball around zero:

$$\mathcal{G}_\epsilon := \{g \in \mathcal{G} : \max\{M(g), L(g)\} \leq \epsilon\}.$$

$[x_1, p_1; x_2, p_2, \dots; x_n, p_n]$  represents a gamble which takes values  $x_1, x_2, \dots, x_n$  with respective probabilities of  $p_1, p_2, \dots, p_n$ .<sup>15</sup>

An *index of riskiness* is a function  $Q : \mathcal{G} \rightarrow \mathbb{R}_+$  which associates each gamble with a positive real. Note that an index of riskiness is *objective*, in the sense that its value depends only on the gamble and not on any agent-specific attribute. An index of riskiness  $Q$  is *homogeneous* (of degree  $k$ ) if  $Q(tg) = t^k \cdot Q(g)$  for all  $t > 0$  and all gambles  $g \in \mathcal{G}$ .

$Q^{AS}(g)$ , the *Aumann-Serrano index of riskiness* of gamble  $g$ , is implicitly defined by the equation

$$\mathbb{E} \left[ \exp \left( -\frac{g}{Q^{AS}(g)} \right) \right] = 1.$$

$Q^{FH}(g)$ , the *Foster-Hart measure of riskiness* of  $g$ ,<sup>16</sup> is implicitly defined by the equation

---

<sup>15</sup>This notation will not be used when it is important to distinguish between random variables and distributions.

<sup>16</sup>I also refer to  $Q^{FH}$  as an index of riskiness.

$$\mathbb{E} \left[ \log \left( 1 + \frac{g}{Q^{FH}(g)} \right) \right] = 0.$$

Note that both  $Q^{AS}$  and  $Q^{FH}$  are homogeneous of degree 1. Additionally, these indices are monotone with respect to first and second order stochastic dominance;<sup>17</sup> namely, if  $g$  is stochastically dominated by  $g'$  then  $Q^{AS}(g) > Q^{AS}(g')$  and also  $Q^{FH}(g) > Q^{FH}(g')$  (Aumann and Serrano, 2008; Foster and Hart, 2009).

*Value at Risk* (VaR) is a family of indices commonly used in the financial industry (Artzner, 1999; Aumann and Serrano, 2008). VaR indices depend on a parameter called the *confidence level*. For example, the VaR of a gamble at the 95 percent confidence level is the largest loss that occurs with probability greater than 5 percent.

In this paper, a *utility function* is a von Neumann–Morgenstern utility function for money. I assume that utility functions are strictly increasing, strictly concave and twice continuously differentiable unless otherwise mentioned. The *Arrow-Pratt coefficient of absolute risk aversion* (ARA),  $\rho$ , of  $u$  at wealth  $w$  is defined

$$\rho_u(w) := -\frac{u''(w)}{u'(w)}.$$

The *Arrow-Pratt coefficient of relative risk aversion* (RRA),  $\varrho$ , of  $u$  at wealth  $w$  is defined

$$\varrho_u(w) := -w \frac{u''(w)}{u'(w)}.$$

---

<sup>17</sup>A gamble  $g$  *first order stochastically dominates*  $h$  iff for every weakly increasing (not necessarily concave) utility function  $u$  and every  $w \in \mathbb{R}$ ,  $\mathbb{E}[u(w+g)] \geq \mathbb{E}[u(w+h)]$ , with strict inequality for at least one such function. A gamble  $g$  *second order stochastically dominates*  $h$  iff for every weakly concave utility function  $u$  and every  $w \in \mathbb{R}$ ,  $\mathbb{E}[u(w+g)] \geq \mathbb{E}[u(w+h)]$ , with strict inequality for at least one such function.

Note that  $\rho_u(\cdot)$  and  $\varrho_u(\cdot)$  are utility specific attributes and that both  $\rho$  and  $\varrho$  yield a complete order on utility-wealth pairs. That is, the risk aversion, as measured by  $\rho$  (or  $\varrho$ ), of any two agents with two given wealth levels can be compared.

A gamble  $g$  is *accepted* by  $u$  at wealth  $w$  if  $\mathbb{E}[u(w + g)] > u(w)$ , and is *rejected* otherwise. Given an index of riskiness  $Q$ , a utility function  $u$ , a wealth level  $w$  and  $\epsilon > 0$ :

**Definition.**  $R_Q^\epsilon(u, w) := \sup \{Q(g) | g \in \mathcal{G}_\epsilon \text{ and } g \text{ is accepted by } u \text{ at } w\}$

**Definition.**  $S_Q^\epsilon(u, w) := \inf \{Q(g) | g \in \mathcal{G}_\epsilon \text{ and } g \text{ is rejected by } u \text{ at } w\}$

$R_Q^\epsilon(u, w)$  is the  $Q$ -riskiness of the *riskiest* accepted gamble according to  $Q$ , restricting the support of the gambles to an  $\epsilon$ -ball.  $S_Q^\epsilon(u, w)$  is the  $Q$ -riskiness of the *safest* rejected gamble according to  $Q$ , again restricting the support of the gambles to an  $\epsilon$ -ball.

**Definition.**  $u$  at  $w$  is (locally) at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  if for every  $\delta > 0$  there exists  $\epsilon > 0$  such that  $S_Q^\epsilon(v, w') \geq R_Q^\epsilon(u, w) - \delta$ .

The interpretation of  $u$  at  $w$  being at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  is that, at least for small gambles, if  $u$  at  $w$  accepts *any* small gamble with a certain level of  $Q$ -riskiness,  $v$  at  $w'$  accepts *all* small gambles which are significantly (by at least  $\delta$ ) less  $Q$ -risky. Alternatively, if  $v$  at  $w'$  rejects *any* small gamble with a certain level of  $Q$ -riskiness,  $u$  at  $w$  rejects *all* small gambles which are significantly (by at least  $\delta$ )  $Q$ -riskier.

The following definitions will also prove useful:

**Definition.**  $R_Q(u, w) := \lim_{\epsilon \rightarrow 0^+} R_Q^\epsilon(u, w)$

**Definition.**  $S_Q(u, w) := \lim_{\epsilon \rightarrow 0^+} S_Q^\epsilon(u, w)$ .<sup>18</sup>

Roughly speaking,  $R_Q(u, w)$  is the  $Q$ -riskiness of the  $Q$ -riskiest “local gamble” that  $u$  accepts at  $w$ , and  $S_Q(u, w)$  is the  $Q$ -riskiness of the  $Q$ -safest “local gamble” that is rejected by  $u$  at  $w$ . The inverse of  $R_Q$  and  $S_Q$  is a natural measure of the aversion to  $Q$ -riskiness.<sup>19</sup> The reason is that  $R_Q$  is high for utility-wealth pairs in which  $Q$ -risky gambles are accepted, so a reasonable  $Q$ -riskiness aversion measure should imply that the aversion to  $Q$ -riskiness at such utility-wealth is low. Similarly,  $S_Q$  is low at a given utility-wealth pair when  $Q$ -safe gambles are rejected, so the measure of local aversion to  $Q$ -riskiness must be high in this case.

The *coefficient of local aversion to  $Q$ -riskiness of  $u$  at  $w$*  is therefore defined as

$$A_Q(u, w) := \frac{1}{R_Q(u, w)},$$

noting that unless otherwise mentioned, all of the results would hold for  $\frac{1}{S_Q(u, w)}$  as well. As is shown below, this definition makes it possible to discuss the ordinal equivalence of the coefficient of local aversion to  $Q$ -riskiness, which depends both on agents behavior and on the properties of the index  $Q$ , with orders such as ARA or RRA, which depend on the preferences exclusively, and are independent of the index.

---

<sup>18</sup>The existence of the limit in the wide sense is guaranteed by the fact that the suprema (infima) in the definitions of  $R^\epsilon$  ( $S^\epsilon$ ) are taken on nested supports.

<sup>19</sup>For our purposes,  $0 = \infty^{-1}$  and  $\infty = 0^{-1}$ .

### 1.3 Local Aversion to $Q$ -Riskiness

Since no restrictions on  $Q$  were made (other than possibly homogeneity), at this point coefficients of local aversion to  $Q$ -riskiness might look like a class of arbitrary orderings over  $(u, w)$  pairs. However, I claim that its members are connected to the standard concepts of local risk aversion. One reason is that they induce orderings which *refine* the following *natural partial order* (Yaari, 1969):  $u$  at  $w$  is locally no less risk averse than  $v$  at  $w'$  (written  $(u, w) \succ (v, w')$ ) if and only if there exists  $\epsilon > 0$  such that for every  $g \in \mathcal{G}_\epsilon$ , if  $u$  accepts  $g$  at  $w$  then so does  $v$  at  $w'$ . An order  $O$  refines the natural partial order if for all  $g$  and  $h$ ,  $g \succ h \implies gOh$ .

**Lemma 1.1.** For every index of riskiness  $Q$ , the order induced by  $A_Q$  refines the natural partial order.

*Proof.* Assume that  $(u, w) \succ (v, w')$ . Then there exists  $\epsilon' > 0$  such that for every  $g \in \mathcal{G}_{\epsilon'}$  if  $u$  accepts  $g$  at  $w$  then so does  $v$  at  $w'$ . As in the definition of  $R_Q$  we have  $\epsilon \rightarrow 0^+$ , disregarding all  $\epsilon \geq \epsilon'$  will not change the result. Note that for every  $\epsilon < \epsilon'$

$$\{Q(g) \mid g \in \mathcal{G}_\epsilon \text{ and } g \text{ is accepted by } u \text{ at } w\} \subseteq$$

$$\{Q(g) \mid g \in \mathcal{G}_\epsilon \text{ and } g \text{ is accepted by } v \text{ at } w'\}.$$

This means that for every  $\epsilon < \epsilon'$ ,  $R^\epsilon(u, w) \leq R^\epsilon(v, w')$  as the suprema in the definition of  $R_Q^\epsilon(v, w')$  are taken on a superset of the corresponding sets in the definition of  $R_Q^\epsilon(u, w)$ . The result follows as weak inequalities are preserved in the limit.  $\square$

Next, I show that the coefficient of local aversion to AS (FH) riskiness gives rise to a complete order which coincides with the one implied by the Arrow-Pratt ARA coefficient.

**Lemma 1.2.** For every utility function  $u$  and every  $w$ ,  $R_{QAS}(u, w) = S_{QAS}(u, w)$  and  $A_{QAS}(u, w) = \rho_u(w)$ .

*Proof.* First, observe that if  $u$  and  $v$  are two utility functions and there exists an interval  $I \subseteq \mathbb{R}$  such that  $\rho_u(x) \geq \rho_v(x)$  for every  $x \in I$ , then for every wealth level  $w$  and lottery  $g$  such that  $w + g \subset I$ , if  $g$  is rejected by  $v$  at  $w$  it is also rejected by  $u$  for the same wealth level. Put differently, if  $g$  is accepted by  $u$  at  $w$  it is also accepted by  $v$  at the same wealth level. The reason is that the condition implies that in this domain,  $u$  is a concave transformation of  $v$  (Pratt, 1964), hence by Jensen's inequality  $u(w) \leq \mathbb{E}[u(w + g)]$  implies that  $v(w) \leq \mathbb{E}[v(w + g)]$ .

Keeping in mind that  $u'(x) > 0$ , we have that  $\rho_u(x)$  is continuous. Specifically,

$$\forall \delta > 0 \exists \epsilon > 0 \text{ s.t. } x \in (w - \epsilon, w + \epsilon) \Rightarrow |\rho_u(x) - \rho_u(w)| < \delta. \quad (1.3.0.1)$$

Recall that a CARA utility function with ARA coefficient of  $\alpha$  rejects all gambles with AS riskiness greater than  $\frac{1}{\alpha}$  and accepts all gambles with AS riskiness smaller than  $\frac{1}{\alpha}$  (Aumann and Serrano, 2008). For any  $\delta < \rho_u(w)$ , given an  $\epsilon$ -environment of  $w$  in which  $\rho_u \in (\rho_u(w) - \delta, \rho_u(w) + \delta)$ , taking the CARA functions with ARA of  $\rho_u(w) + \delta$  and  $\rho_u(w) - \delta$ , and applying the first observation (where  $I$  is  $(w - \epsilon, w + \epsilon)$ ) completes the proof.  $\square$

Lemma 1.2 essentially shows that every utility function may be approxi-



mated locally using CARA functions, which are well-behaved with respect to the AS index. Given the ARA of  $u$  at a given wealth level, I take two CARA utility functions, one with slightly higher ARA, and the other with slightly lower ARA. For small environments around the given wealth level,  $\rho_u$  is almost constant, so the two CARA functions “sandwich” the utility function in terms of ARA. This implies that for small gambles, one CARA function accepts more gambles than  $u$ , and the other less gambles, in the sense of set inclusion. Since CARA functions accept and reject exactly according to an AS riskiness cutoff, and since cutoffs are close for similar ARA values, it follows that the coefficient of local aversion to AS-riskiness is pinned down completely.

**Lemma 1.3.** For every utility function  $u$  and every  $w$ ,  $R_{Q^{FH}}(u, w) = S_{Q^{FH}}(u, w)$  and  $A_{Q^{FH}}(u, w) = \rho_u(w)$ .

*Proof.* According to Statement 4 in Foster and Hart (2009):

$$-L(g) \leq Q^{AS}(g) - Q^{FH}(g) \leq M(g). \quad (1.3.0.2)$$

Therefore, if  $g \in \mathcal{G}_\epsilon$  then:

$$|Q^{AS}(g) - Q^{FH}(g)| \leq \epsilon. \quad (1.3.0.3)$$

From Inequality 1.3.0.3 one can deduce that  $R_{Q^{FH}}(u, w) = R_{Q^{AS}}(u, w)$  and  $S_{Q^{FH}}(u, w) = S_{Q^{AS}}(u, w)$ . Lemma 1.2 completes the proof.  $\square$

The result of Lemma 1.3 is not surprising in light of Lemma 1.2, as Foster and Hart (2009) already noted that the Taylor expansions around 0 of the functions that define  $Q^{FH}$  and  $Q^{AS}$  differ only from the third term on.

Roughly speaking, this means that for gambles with small supports  $Q^{AS}$  and  $Q^{FH}$  are close.

Theorem 1.1 summarizes the results of Lemmata 1.1-1.3.

**Theorem 1.1.** (i) *For any index of riskiness  $Q$ ,  $A_Q$  refines the natural partial order.* (ii) *For every utility function  $u$  and every  $w$ ,  $A_{Q^{AS}}(u, w) = A_{Q^{FH}}(u, w) = \rho_u(w)$ . Furthermore,  $R_{Q^{AS}}(u, w) = S_{Q^{AS}}(u, w)$  and  $R_{Q^{FH}}(u, w) = S_{Q^{FH}}(u, w)$ .*

**Corollary 1.1.** *For  $Q \in \{Q^{AS}, Q^{FH}\}$   $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  iff  $\rho_u(w) \geq \rho_v(w')$ .*

Note that part (i) of Theorem 1.1 states that the order induced by  $A_Q$  refines the weak, no-less risk averse, partial order, and not the strict one. The strict version of this statement is not correct as the following example demonstrates. The example also shows that it is not the case that for all popular risk indices the coefficient of local aversion is equal to  $\rho$  or refines the order it induces, and that the same is true for the relation at least as averse to  $Q$ -riskiness.

**Example 1.1.** For any confidence level  $\alpha \in (0, 1)$ , for all agents and wealth levels, the coefficient of local aversion to  $Q(\cdot) := \exp\{\text{VaR}_\alpha(\cdot)\}$  is equal to 1, and any agent at any wealth level is at least as averse to  $Q$ -riskiness as any other agent.<sup>20</sup>

It is noteworthy that the example would go through with the exponent of any *coherent risk measure* (Artzner *et al.*, 1999). The fact that these

---

<sup>20</sup>The exponent is only used to assure that the index is positive. It has no ordinal effect.

indices are not well suited for the task of comparing agents' preferences is not surprising. These indices are motivated by the problem of setting a minimal reserve requirements for investors in a given position (Artzner, 1999), and so they take the point of view of a regulator, not the investor.

Up until this point, I showed that the local aversion to AS and FH riskiness induces the same order as the ARA coefficient, the standard measure of local risk aversion, and that the coefficient of local aversion to AS and FH riskiness is in fact equal to the ARA coefficient. This means that one can start with a small set of axioms, namely Aumann and Serrano's (2008) or Foster and Hart's (2013), and define a complete order of riskiness over gambles. Then, the coefficient of local aversion of agents to riskiness can be derived, and it will be equal to the well-known Arrow-Pratt coefficient. The relation at least as averse to AS (FH)-riskiness will also induce the same order. Hence, both AS and FH satisfy the desirable property that less risk averse agents according to ARA accept riskier gambles according to AS or FH.

Theorem 1.1 and Corollary 1.1 might be interpreted as evidence that AS and FH were "well-chosen" in some sense. However, I will show in Theorem 1.2 that while AS and FH satisfy the desirable properties mentioned above, there are other indices which satisfy the same properties. Moreover, some of these indices are not "reasonable" in the sense that they are not monotone with respect to first order stochastic dominance, in clear violation of the requirement that an index of riskiness should judge as riskier the alternative risk-averse individuals less prefer. Theorem 1.3 further identifies sufficient conditions on  $Q$  under which the coefficient of local aversion to  $Q$ -riskiness and the relation at least as averse to  $Q$ -riskiness yield the same order as the

Arrow-Pratt (local) absolute risk aversion. Before stating these results, I must introduce a few key properties.

**Axiom** (Homogeneity).  *$Q$  is homogeneous of degree  $k$  for some  $k > 0$ .*

The homogeneity axiom has both cardinal and ordinal content. For the case  $k = 1$ , its cardinal interpretation is that doubling the stakes doubles the riskiness. The ordinal content is that doubling the stakes increases the riskiness. When taking the point of view of an agent, not a regulator setting a minimal reserve requirement, the cardinal part is not necessarily desirable. In what follows, I assume it for its simplicity and since homogeneity of degree 1 appears in the original axiomatic characterization of the AS index, but later I remove this axiom.

**Axiom** (Local consistency).  $\forall u \forall w \exists \lambda > 0 \forall \delta > 0 \exists \epsilon > 0 R_Q^\epsilon(u, w) - \delta < \lambda < S_Q^\epsilon(u, w) + \delta$ .

Local consistency says that small gambles that are significantly  $Q$ -safer than some cut-off level are always accepted, and that ones significantly riskier than the cutoff are always rejected. Lemma A.2 in the appendix shows that whenever homogeneity is satisfied, local consistency implies that  $0 < S_Q(u, w) = R_Q(u, w) < \infty$ . This means, that for “small” gambles  $Q$  is sufficient information to determine an agent’s optimal behavior. In other words, the decisions of agents are consistent with the index, on small domains.

**Definition** (Reflexivity). The relation at least as averse to  $Q$ -riskiness is *reflexive* if for all  $u$  and  $w$ ,  $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $u$  at  $w$ .

**Proposition.** *If  $Q$  satisfies local consistency, then the relation locally at least as averse to  $Q$ -riskiness is reflexive.*

**Definition** (Ordinally equivalent). Given an index of riskiness  $Q$ ,  $A_Q$  is *ordinally equivalent* to the coefficient of absolute risk aversion  $\rho$ , if  $\forall u, v \ \forall w, w'$   
 $A_Q(u, w) > A_Q(v, w') \iff \rho_u(w) > \rho_v(w')$ .

**Theorem 1.2.** (i) There exists a continuum of locally consistent, homogeneous of degree 1, riskiness indices for which the coefficient of local aversion equals the Arrow-Pratt coefficient.<sup>21</sup> (ii) Moreover, some of these indices are not monotone with respect to first order stochastic dominance.

(i) is proved in the appendix using the observation that for every  $a > 0$  any combination of the form  $Q_a(\cdot) := Q^{FH}(\cdot) + a \cdot |Q^{FH}(\cdot) - Q^{AS}(\cdot)|$  is an index of riskiness for which the coefficient of local aversion equals the coefficient of local aversion to  $Q^{FH}$ . The reason this holds is that for small supports, the second element in the definition is vanishingly small by Inequality 1.3.0.3, and so  $Q_a$  and  $Q^{FH}$  should be close. The following example demonstrates (ii).

**Example 1.2.** Take  $Q_1(\cdot) := Q^{FH}(\cdot) + |Q^{FH}(\cdot) - Q^{AS}(\cdot)|$  and  $g = [1, \frac{\epsilon}{1+\epsilon}; -1, \frac{1}{1+\epsilon}]$ .  $Q^{AS}(g) = 1$  and  $Q^{FH}(g) \approx 1.26$ , hence  $Q_1(g) < 1.6$ . Now take  $g' = [1, 1 - \epsilon; -1, \epsilon]$ . For small values of  $\epsilon$ ,  $Q^{AS}(g') \approx 0$  but  $Q^{FH}(g') > 1$ , so  $Q_1(g') > 1.6$ . Therefore, while  $g'$  first order stochastically dominates  $g$ ,  $Q_1(g) < Q_1(g')$ .

---

<sup>21</sup>Omitting the homogeneity of degree 1 requirement would yield a trivial statement as, for example, an arbitrary change of the values of  $Q^{AS}$  for gambles taking values larger than some  $M > 0$  will result in a valid index. The requirement that the local aversion to the index coincides with the Arrow-Pratt coefficient, and not just with the order it implies, is a normalization that rules out, for example, the use of positive multiples of  $Q^{AS}$ .

**Theorem 1.3.** *If  $Q$  satisfies local consistency and homogeneity of degree  $k > 0$ , then  $A_Q$  is ordinally equivalent to  $\rho$ , and the relation at least as averse to  $Q$ -riskiness induces the same order as  $\rho$ .<sup>22</sup>*

The proof is in the appendix. It extends the reasoning of Lemma 1.1.

*Remark.* Both axioms in Theorem 1.3 are essential: As the following examples demonstrate, omitting either admits indices for which the coefficient of local aversion is not ordinally equivalent to  $\rho$ , and the relation at least as averse to  $Q$ -riskiness does not induce the same order as  $\rho$ .

**Example 1.3.**  $Q(\cdot) \equiv 5$  satisfies local consistency, but it does not satisfy homogeneity of degree  $k > 0$ . The local aversion to this index induces the trivial order and  $A_Q \equiv \frac{1}{5}$ .

**Example 1.4.**  $Q(\cdot) = \mathbb{E}[\cdot]$  is homogeneous of degree 1, but it violates local consistency. The local aversion to this index induces the trivial order and  $A_Q \equiv \infty$ .

Later in the paper, homogeneity will sometimes no longer be required. It will be replaced by a requirement of continuity and monotonicity with respect to first order stochastic dominance (or mean preserving spreads). For completeness, I present an example of a locally consistent index which satisfies continuity and monotonicity with respect to first and second order stochastic dominance but does not possess the ordinal content of homogeneity.

**Definition (Continuity).** An index of riskiness  $Q$  is continuous if  $Q(g) =$

---

<sup>22</sup>To be precise, this statement means that  $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  if and only if  $\rho_u(w) \geq \rho_v(w')$ .

$\lim_{n \rightarrow \infty} Q(g_n)$  whenever  $g_n$  are uniformly bounded gambles which converge to  $g$  in probability.

**Example 1.5.**  $Q(\cdot) = \exp\{Q^{AS}(\cdot) - \mathbb{E}[\cdot]\}$  inherits its positivity from the exponent, it is continuous and monotonic with respect to first order stochastic dominance as both  $Q^{AS}(\cdot)$  and  $-\mathbb{E}[\cdot]$  are. It is monotonic with respect to second order stochastic dominance as  $Q^{AS}(\cdot)$  is increasing in mean-preserving spreads and  $\mathbb{E}$  does not vary with mean-preserving spreads (weakly increasing).  $Q$  satisfies local consistency as for small supports it is almost equal to  $\exp\{Q^{AS}(\cdot)\}$ , which is locally consistent. Finally, for  $g$  such that  $Q^{AS}(g) < \mathbb{E}[g]$  and  $\lambda > 1$ ,  $Q(\lambda g) < Q(g)$ . For small  $\epsilon > 0$ , gambles of the form  $g = [-\epsilon, \frac{1}{2}; 1, \frac{1}{2}]$  satisfy the required inequality.

## 1.4 The Aversion to AS-Riskiness and the Demand for Gambles

Samuelson (1960) shows that *“if you would always refuse to take favorable odds on a single toss, you must rationally refuse to participate in any (finite) sequence of such tosses”* (Samuelson, 1963). But Samuelson (1963) also warns against undue extrapolation of his theorem saying *“It does not say that one must always refuse a sequence if one refuses a single venture: if, at higher income levels the single losses become acceptable, and at lower levels the penalty of losses does not become infinite, there might well be a long sequence that it is optimal.”* The following proposition shows that AS has properties which generalize the property discussed by Samuelson.

**Proposition 1.1.** *A gamble  $g$  with  $Q^{AS}(g) = c$  is rejected by  $u$  at  $w$  only if there exist some  $w' \in [w - L(g), w + M(g)]$  such that small gambles with  $Q^{AS}$  of  $c$  are rejected at  $w'$ . A gamble  $g$  with  $Q^{AS}(g) = c$  is accepted by  $u$  at  $w$  only if there exist some  $w' \in [w - L(g), w + M(g)]$  such that small gambles with  $Q^{AS}$  of  $c$  are accepted at  $w'$ .*

*Proof.* Omitted. □

**Corollary 1.2.** *If  $Q^{AS}(g) > \sup_w \{A_{Q^{AS}}^{-1}(u, w)\} = \sup_w \{\rho_u^{-1}(w)\}$  then  $u$  rejects  $g$  at any wealth level. If  $Q^{AS}(g) < \inf_w \{A_{Q^{AS}}^{-1}(u, w)\} = \inf_w \{\rho_u^{-1}(w)\}$  then  $u$  accepts  $g$  at any wealth level.*

The corollary suggests a partition of the class of gambles into three: “risky” gambles, which the agent never accepts, “safe” gambles which are always accepted, and gambles whose acceptance is subject to wealth effects. Knowing the distribution of preferences in a given population, the intersection of the relevant “risky” and “safe” segments yields a partition which is mutually agreed upon. Such a partition could be used as a simple tool for evaluating policies, as I will show in Section 1.6. It may also be used as a simple tool for providing bounds on risk attitudes, as illustrated in the following example.

**Example 1.6.** Say that a population of agents are observed making acceptance and rejection decisions on gambles. Say that  $A$  is the set of gambles rejected by some agent, and  $B$  is the set of gambles accepted by some agent. Then if, for some  $g \in B$  and for all  $u$ ,  $Q^{AS}(g) > \sup_w \{\rho_u^{-1}(w)\}$ , a contradiction would be implied. So, for some  $u$ ,  $\max_{g \in B} Q^{AS}(g) \leq \sup_w \{\rho_u^{-1}(w)\}$ . Similarly  $\min_{g \in A} Q^{AS}(g) \geq \inf_w \{\rho_v^{-1}(w)\}$  for some  $v$ .



The next result shows a property of the index which is in the spirit of Samuelson's argument, and in fact implies Samuelson's theorem. It shows that the sets of "risky" and "safe" gambles are closed under compounding of independent gambles.

**Definition** (Compound gamble property). An index  $Q$  has the *compound gamble* property if for every compound gamble of the form  $f = g + \mathbf{1}_A h$ , where  $\mathbf{1}$  is an indicator,  $A$  is an event such that  $g$  is constant on  $A$  ( $g|_A \equiv x$  for some  $x$ ) and  $h$  is independent of  $A$ ,  $\max \{Q(g), Q(h)\} \geq Q(f) \geq \min \{Q(g), Q(h)\}$ .

**Proposition 1.2.**  $Q^{AS}$  satisfies the compound gamble property. Thus, if  $g, h \in \mathcal{G}$  are independent, and  $\min \{Q^{AS}(g), Q^{AS}(h)\} > \sup_w \{\rho_u^{-1}(w)\}$ , then a compound gamble of  $g$  and  $h$  will also satisfy the inequality. Additionally, if  $g, h \in \mathcal{G}$  are independent, and  $\max \{Q^{AS}(g), Q^{AS}(h)\} < \inf_w \{\rho_u^{-1}(w)\}$ , then a compound gamble of  $g$  and  $h$  will also satisfy the inequality.

*Proof.* See appendix. □

Theorem 1.3 identifies conditions under which the coefficient of local aversion to  $Q$ -riskiness and the relation at least as averse to  $Q$ -riskiness induce the same order as the Arrow-Pratt ARA. But according to Theorem 1.2 and Example 1.5 this property is not enough to characterize a "reasonable" index of riskiness. These findings call for additional requirements from an index of riskiness. I propose a *generalized Samuelson property* and show that with reflexivity (local consistency), monotonicity and continuity it pins down uniquely the AS index.

**Axiom** (Generalized Samuelson property).  $\forall u, w' \ S_Q^\infty(u, w') \geq \inf_w S_Q(u, w)$   
and  $R_Q^\infty(u, w') \leq \sup_w R_Q(u, w)$ .

The axiom says that no agent accepts a large gamble of a certain degree of riskiness if he rejects small ones of the same degree of riskiness at any wealth level, and no agent rejects a large gamble of a certain degree of riskiness if he accepts small ones of the same degree of riskiness at any wealth level.

**Theorem 1.4.**  $(Q^{AS})^k$  is the unique index of riskiness that satisfies local consistency, global consistency and homogeneity of degree  $k > 0$ , up to a multiplication by a positive number.

*Proof.* See appendix. □

As was discussed previously, the cardinal content of the homogeneity axiom is not necessarily appealing for general indices of riskiness. In what follows, this axiom will be removed and replaced with less demanding conditions: monotonicity with respect to first order stochastic dominance and continuity. The combination of the generalized Samuelson property, monotonicity, continuity and reflexivity of the relation locally at least as averse to  $Q$ -riskiness implies local consistency, and so the local consistency requirement could be replaced with the weaker requirement of reflexivity.

**Theorem 1.5.** If  $Q$  satisfies the generalized Samuelson property, reflexivity, monotonicity with respect to first order stochastic dominance and continuity, then  $Q$  is ordinally equivalent to  $Q^{AS}$ .

*Proof.* See appendix. □

*Claim 1.1.* The monotonicity requirement in the theorem could be replaced by each of the following conditions:

- (a) Monotonicity with respect to mean-preserving spreads
- (b) Satisfying the ordinal content of homogeneity
- (c) Monotonicity with respect to increases in the lowest value of the gamble, leaving the rest of the values unchanged

In such case, monotonicity with respect to first order stochastic dominance will be a result, not an assumption.

**Corollary 1.3.** *The FH index of riskiness does not satisfy the generalized Samuelson property.*

**Example 1.7.**  $g := [1, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}]$  has  $Q^{FH}(g) = 1$ . When compounding 3 i.i.d gambles with this distribution, the largest loss that happens with positive probability is -1.5. This implies that the FH-riskiness of the compound gamble must be at least 1.5.

## 1.5 Global Consistency

The generalized Samuelson property implies that if for two agents,  $u$  and  $v$ ,  $\inf_{w'} S_Q(v, w') \geq \sup_w R_Q(u, w)$ , then  $S_Q^\infty(v, w_0) \geq R_Q^\infty(u, w_1)$  at any two wealth levels  $w_0$ , and  $w_1$ . In this section, I propose *global consistency* – a weaker restriction on pairs of agents. The following definition is required in order to state the condition.

**Definition** (Globally more averse to  $Q$ -riskiness). Let  $Q$  be an index of riskiness.  $u$  is *globally at least as averse to  $Q$ -riskiness as  $v$*  is (written

$u \succeq_Q v$ ) if, for every  $w$  and  $w'$ ,  $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $v$  at  $w'$ .  $u$  is *globally more averse to  $Q$ -riskiness than  $v$*  (written  $u >_Q v$ ) if  $u \succeq_Q v$  and not  $v \succeq_Q u$ .<sup>23</sup>

**Axiom** (Global consistency). *For every pair of utilities  $u$  and  $v$ , for every  $w$  and every  $g$  and  $h$  in  $\mathcal{G}$ , if  $u >_Q v$ ,  $u$  accepts  $g$  at  $w$ , and  $Q(g) > Q(h)$ , then  $v$  accepts  $h$  at  $w$ .*

*Claim 1.2.* Global consistency is implied by the generalized Samuelson property.

The axiom of global consistency is a weak requirement, in the sense that it imposes no restriction for pairs of utilities which cannot be compared using the partial order globally more averse to  $Q$ -riskiness. It is inspired by the duality axiom of AS. For small gambles, it follows immediately from local consistency. In fact, local consistency could have been stated in a very similar way, had it been assumed that the relation at least as averse to  $Q$ -riskiness is reflexive. It would state that if  $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  is, then there exists  $\lambda > 0$  such that for all  $\delta > 0$  there exists  $\epsilon > 0$  with  $R_Q^\epsilon(u, w) - \delta < \lambda < S_Q^\epsilon(v, w') + \delta$ . Roughly, it states that if the risk averse agent accepts a small gamble with a certain level of riskiness, the less risk averse agent will accept small gambles which are  $Q$ -safer. The content of the axiom of global consistency comes from the fact that it places no restriction on the support of gambles, so that when two agents that can be compared by the partial order “globally more averse to  $Q$ -riskiness,” the axiom requires

---

<sup>23</sup>The above definition is different from the AS definition of uniformly more risk-averse. It is derived directly from the index  $Q$  and the utility function  $u$ . However, if the relation at least as averse to  $Q$ -riskiness induces the same order as  $\rho$  the two definitions are equivalent.

that the less averse agent accepts  $Q$ -riskier gambles, and the requirement applies not only for small gambles.

**Theorem 1.6.**  $(Q^{AS})^k$  is the unique index of riskiness that satisfies local consistency, global consistency and homogeneity of degree  $k > 0$ , up to a multiplication by a positive number.

*Proof.* Let  $Q$  be homogeneous of degree 1. From Theorem 1.3,  $A_Q$  is ordinally equivalent to  $\rho$ , and the relation at least as averse to  $Q$ -riskiness induces the same order as  $\rho$ . The AS duality axiom states that if  $u$  is uniformly more averse to risk than  $v$ ,  $u$  accepts  $g$  at  $w$ , and  $Q(g) > Q(h)$ , then  $v$  accepts  $h$  at  $w$ . That the relation at least as averse to  $Q$ -riskiness induces the same order as  $\rho$  means that  $u$  is globally more averse to  $Q$ -riskiness than  $v$  if and only if  $u$  is uniformly more risk averse than  $v$ . With global consistency, this implies the duality axiom. But the only indices that satisfy homogeneity of degree 1 and the duality axiom are positive multiples of  $Q^{AS}$  (Aumann and Serrano, 2008). If  $Q$  is homogeneous of degree  $0 < k \neq 1$ ,  $Q' = (Q)^{\frac{1}{k}}$  is homogeneous of degree 1, and still satisfies the other properties,<sup>24</sup> so  $Q'$  must equal  $C \cdot Q^{AS}$  for some  $C > 0$ , and so  $Q$  is equal to  $C^k \cdot (Q^{AS})^k$ . Finally, Theorems 1.1 and 1.3 and the discussion above imply that for all  $k > 0$ ,  $(Q^{AS})^k$  satisfies the axioms,<sup>25</sup> and the same holds for its positive multiples.  $\square$

**Corollary 1.4.**  $Q^{FH}$ , the FH index of riskiness, does not satisfy global consistency.

---

<sup>24</sup>To verify this, note that  $f(x) = x^{\frac{1}{k}}$  is continuous, and  $Q$  and  $Q'$  are ordinally equivalent.

<sup>25</sup>In fact, this was shown only for the case  $k = 1$ , but it is clear that the other cases are implied by this case.

**Example 1.8.** Consider a gamble  $g = [1, \frac{\epsilon}{1+\epsilon}; -1, \frac{1}{1+\epsilon}]$ ,  $Q^{AS}(g) = 1$  and  $Q^{FH}(g) \approx 1.26$ , and a gamble  $g' = [2, 1 - \epsilon; -2, \epsilon]$ . For small values of  $\epsilon$ ,  $Q^{AS}(g') \approx 0$  but  $Q^{FH}(g') > 2$ . Hence  $Q^{AS}(g) > Q^{AS}(g')$  yet  $Q^{FH}(g) < Q^{FH}(g')$ . Since the local aversion to FH-riskiness is equal to the local aversion to AS-riskiness by Theorem 1.1, any two CARA utility functions with different ARA between  $\frac{1}{Q^{AS}(g)}$  and  $\frac{1}{Q^{AS}(g')}$  together with the two gambles violate global consistency.

In what follows, the homogeneity axiom will be removed and replaced with less demanding conditions: monotonicity with respect to first order stochastic dominance and continuity. Example 1.9 will show that these axioms will not suffice for assuring that the coefficient of local aversion to  $Q$ -riskiness is non-degenerate, or even to ensure that the index is monotonic with respect to second order stochastic dominance, and so I will require a slightly stronger version of global consistency. On the other hand, the combination of strong global consistency, monotonicity, continuity and reflexivity of the relation locally at least as averse to  $Q$ -riskiness implies local consistency, and so the local consistency requirement could be replaced with the weaker requirement of reflexivity.

**Example 1.9.** Let  $Q(\cdot) = \exp\{-\mathbb{E}[\cdot]\}$ . It is positive, continuous, monotonic with respect to first order stochastic dominance and locally consistent. Additionally, every  $u$  is globally at least as averse to  $Q$ -riskiness as any  $v$ . Hence, no agent is globally more averse to  $Q$ -riskiness than another, and so global consistency is satisfied. The coefficient of local aversion to  $Q$ -riskiness is equal to 1 identically. Finally, mean preserving spreads do not change the

value of the index.

**Axiom** (Strong global consistency). *For every pair of utilities  $u$  and  $v$ , for every  $w$  and every  $g$  and  $h$  in  $\mathcal{G}$ , if  $u \succeq_Q v$ ,  $u$  accepts  $g$  at  $w$ , and  $Q(g) > Q(h)$ , then  $v$  accepts  $h$  at  $w$ .*

The difference between the two axioms is that the weak version uses  $>_Q$  while the strong one uses  $\succeq_Q$ . The strong version, therefore, requires more, as it has a bite for more pairs of utilities. Note that this axiom is violated by the index from Example 1.9. To see this, observe that any two agents  $u$  and  $v$  satisfy both  $u \succeq_Q v$  and  $v \succeq_Q u$ , so  $Q$  must be degenerate in order to satisfy the axiom, but it is not.

*Claim 1.3.* Strong global consistency is implied by the generalized Samuelson property.

**Theorem 1.7.** *If  $Q$  is a continuous index of riskiness that satisfies monotonicity with respect to first order stochastic dominance and strong global consistency, and the relation at least as averse to  $Q$ -riskiness is reflexive, then  $Q$  is ordinally equivalent to  $Q^{AS}$ .*

*Proof.* See appendix. □

**Corollary 1.5.** *If  $Q$  is a continuous index of riskiness that satisfies monotonicity with respect to first order stochastic dominance and strong global consistency and the relation at least as averse to  $Q$ -riskiness is reflexive, then  $Q$  satisfies local consistency and  $A_Q$  is ordinally equivalent to  $\rho$ .*

*Claim 1.4.* The monotonicity requirement in the theorem could be replaced by each of the following conditions:

- (a) Monotonicity with respect to mean-preserving spreads
- (b) Satisfying the ordinal content of homogeneity
- (c) Monotonicity with respect to increases in the lowest value of the gamble, leaving the rest of the values unchanged

In such case, monotonicity with respect to first order stochastic dominance will be a result, not an assumption.<sup>26</sup>

## 1.6 A Generalized Sharpe Ratio

This section considers an investor facing the problem of asset allocation between a risk-free asset, with return  $r_f$  and a market portfolio.<sup>27</sup> Fixing  $r_f$ , a *market return*  $r$  is a real-valued random variable such that  $r - r_f \in \mathcal{G}$ . In particular, the *net return*,  $r - r_f$  has a positive expected value and a positive probability to be negative. For each value of  $r_f$ , let  $\mathcal{R}^{r_f}$ , or simply  $\mathcal{R}$  when there is no risk of confusion, denote the class of all such market returns. An index of performance is a collection of functions  $Q_{r_f} : \mathcal{R}^{r_f} \rightarrow \mathbb{R}_+$ , one for each possible value of the risk-free rate.

One well known index of performance is the *Sharpe ratio*, the ratio between the expected net return and its standard deviation.<sup>28</sup> This measure of “risk adjusted returns,” or “reward-to-variability” (Sharpe, 1966), is frequently used as a performance measure for portfolios (Welch, 2008; Kadan and Liu, 2014). Formally, it is defined by:

---

<sup>26</sup>The continuity assumption could also be relaxed, for example, by requiring continuity in payoffs for fixed probabilities.

<sup>27</sup> $r_f$  may be negative but must be greater than  $-1$ .

<sup>28</sup>Note that  $\sigma(r - r_f) \neq 0$  from the assumption that  $r - r_f \in \mathcal{G}$ .



$$Sh_{r_f}(r) = \frac{\mathbb{E}[r - r_f]}{\sigma(r - r_f)}.$$

The validity of this measure relies critically on several assumptions on the distribution of returns as well as on agents' preferences (Meyer, 1987). In particular, for general distributions, the Sharpe ratio is not monotonic with respect to first order stochastic dominance: portfolio  $r_1$  may have returns that are always higher than portfolio  $r_2$  and yet it will be ranked lower according to the index. This normatively undesirable property of the Sharpe ratio is illustrated by the following example, which is based on an example from Aumann and Serrano (2008):

**Example 1.10.** Let  $r_1 = [-1, .02; 1, .98]$ ,  $r_2 = [-1, .02; 1, .49; 2, .49]$  and  $r_f = 0$ .

$$\mathbb{E}[r_1 - r_f] = .96, \quad \sigma(r_1 - r_f) = .28,$$

hence,

$$Sh_{r_f}(r_1) = \frac{.96}{.28} \approx 3.43.$$

But,

$$\mathbb{E}[r_2 - r_f] = 1.45, \quad \sigma(r_2 - r_f) = \frac{7\sqrt{3}}{20},$$

hence,

$$Sh_{r_f}(r_2) = \frac{1.45 \times 20}{7\sqrt{3}} \approx 2.39.$$

The result will continue to hold if we add some small  $\epsilon > 0$  to all of the payoffs of  $r_2$ .

This undesirable property of the Sharpe ratio is related to the fact that it depends only on the first two moments of the distribution. These moments

are sufficient statistic for a normal distribution, and therefore basing an index on them solely may be reasonable under the assumption of normally distributed returns. This assumption is, however, often rejected in settings where the Sharpe ratio is often used (e.g. Fama, 1965; Agarwal and Naik, 2000; Kat and Brooks, 2001). Moreover, a large body of literature documents the importance of higher order moments for investment decisions (e.g. Kraus and Litzenberger, 1976; Kane, 1982; Harvey and Siddique, 2000; Barro, 2006, 2007; Gabaix, 2008).

Recognizing these limitations of the Sharpe ratio as a measure of performance, Kadan and Liu (2014) propose a reinterpretation of the inverse of the AS index of riskiness as a performance measure and show that it may be more favorable than the Sharpe ratio in an empirical setting. Homm and Pigorsch (2012a) propose a different index, which was mentioned originally in AS: the expected net return divided by the AS riskiness. The index is not derived from first principles but is motivated by a “reward-to-risk” reasoning, where the AS riskiness takes the place of  $\sigma$  in the Sharpe ratio. This section asks which of these indices, if any, does the consistency-motivated approach suggest?

The findings of this section support the latter alternative, which coincides with the Sharpe ratio on the domain of normally distributed returns. The index possesses other desirable properties, importantly monotonicity with respect to stochastic dominance and with respect to *stochastic dominance in the presence of a risk-free asset* (Levy and Kroll, 1978),<sup>29</sup> of the first and

---

<sup>29</sup> $r_1$  first (second) order *stochastically dominates*  $r_2$  *in the presence of a risk-free asset*  $r_f$  if for every  $\alpha \geq 0$  there exists  $\beta \geq 0$  such that  $\alpha r_2 + (1 - \alpha) r_f$  is first (second) order stochastically dominated by  $\beta r_1 + (1 - \beta) r_f$ .

second degree.

### 1.6.1 Preliminaries

**Definition.** A *market transaction* is a pair,  $(q, r) \in \mathbb{R}_+ \times \mathcal{R}$ . Denote by  $\mathcal{T}$  the class of all market transactions.

Say that an agent with utility function  $u$  and initial wealth  $w$  accepts a market transaction if

$$\mathbb{E} \left[ u \left( (w - q)(1 + r_f) + q(1 + r) \right) \right] > u(w(1 + r_f)),$$

and rejects it otherwise.

I assume that it is only the net return that matters for the index. That is, by shifting  $r_f$  and all the possible values of  $r$  by a constant, the performance does not change. This is a standard assumption which makes it possible to compare market returns under different risk-free rates. All the results will continue to hold without this assumption, fixing  $r_f$ .

**Axiom** (Translation invariance).  $\forall \lambda > 0, \forall r_f > -1, \forall r \in \mathcal{R}^{r_f} \quad Q_{r_f + \lambda}(r + \lambda) = Q_{r_f}(r)$ .<sup>30</sup>

The next axiom could be interpreted as saying that if the price of a unit of the market portfolio decreases but it continues to yield the same proceeds, then the market performs better. This intuitive notion is the ordinal content of the axiom T of Artzner *et al.* (1999).

**Axiom** (Monotonicity).  $\forall r_f > -1, \forall r \in \mathcal{R}^{r_f}, \forall \lambda > 0, \text{ if } r_f + \lambda \in \mathcal{R}^{r_f} \text{ then } Q_{r_f}(r + \lambda) > Q_{r_f}(r)$ .

---

<sup>30</sup>If  $r$  is in  $\mathcal{R}^{r_f}$  then  $r + \lambda$  is in  $\mathcal{R}^{r_f + \lambda}$ .

With translation invariance, monotonicity is equivalent to the requirement that the same market return should be considered as better performing in the face of a lower risk-free rate.

To motivate the next axiom, assume for a moment that the risk-free rate is 0, and that agents are free to allocate their resources between the market and a risk-free asset. A reasonable requirement is that an index of performance be homogeneous of degree 0, since any portfolio that could be achieved with market return  $r$  could be mimicked when the return is  $\lambda g$  for any  $\lambda > 0$  by scaling the amount of investment by  $\frac{1}{\lambda}$ . This reasoning clearly extends to the net return,  $r - r_f$ , for any  $r_f$  and  $r$ .

**Axiom** (Homogeneity).  $\forall \lambda > 0, \forall r_f > -1, \forall r \in \mathcal{R}^{r_f}, Q_{r_f}(\lambda \cdot (r - r_f) + r_f) = Q_{r_f}(r)$ .

The Sharpe ratio is an example for a performance index that satisfies this property. Unlike in the other settings presented in this paper, the homogeneity axiom here is *ordinal* and has no cardinal implications.

*Claim 1.5.* A continuous index which satisfies translation invariance and monotonicity but fails to satisfy homogeneity of degree 0 is not monotonic with respect to stochastic dominance in the presence of a risk-free asset.<sup>31</sup>

*Proof.* For some  $\lambda > 0$ , say  $Q_{r_f}(\lambda \cdot (r - r_f) + r_f) > Q_{r_f}(r)$ . From translation invariance,  $Q_0(\lambda \cdot (r - r_f)) > Q_0(r - r_f)$ . From continuity, it will also be the case that  $Q_0(\lambda \cdot (r - r_f)) > Q_0(r - r_f + \epsilon)$  for some small  $\epsilon > 0$ . But  $r - r_f + \epsilon$  first order stochastically dominates  $\lambda \cdot (r - r_f)$  in the presence of a risk-free

---

<sup>31</sup>A precise definition of continuity appears later in this section.

asset with 0 rate of return, as discussed in the argument motivating the homogeneity axiom.  $\square$

**Corollary 1.6.** *The index of performance used by Kadan and Liu (2014) violates monotonicity with respect to stochastic dominance in the presence of a risk-free asset.*

**Example 1.11.** Let  $r$  be a market return with  $\mathbb{E}[r] = 1$  and let  $r_f = 0$ . The index proposed by Kadan and Liu (2014) equals to  $\frac{1}{Q^{AS}(r)} > 0$ . Their index for  $\frac{1}{2}r$ , under the same conditions, is  $\frac{2}{Q^{AS}(r)}$ . From the continuity of their index, this implies that for small  $\epsilon > 0$ ,  $\frac{1}{2}r - \epsilon$  performs better than  $r$  in the Kadan-Liu sense.

For  $c \geq 0$  and  $r_f > -1$ , define  $\mathcal{R}_c^{r_f} := \{r \in \mathcal{R}^{r_f} \mid \mathbb{E}[r] = r_f + c\}$ , the class of market returns with expected net return of  $c$ . If  $Q$  satisfies homogeneity, it is completely characterized by the restriction of  $Q_{r_f}$  to  $\mathcal{R}_{r_f+1}$ . If  $Q$  further satisfies translation invariance, then there is no loss of generality in writing  $Q(r - r_f) := Q_0(r - r_f) = Q_{r_f}(r)$ . This means that it is sufficient to consider the case that  $r_f = 0$  and to characterize  $Q : \mathcal{R}_1 \rightarrow \mathbb{R}_+$ . From this point on, unless specifically mentioned, attention will be restricted to this case.

Denote the class of “local” market transactions by

$$\mathcal{T}_\epsilon := \{(q, r) \in \mathcal{T} \mid \max\{qr\} - \min\{qr\} < \epsilon, r \in \mathcal{R}_1\}.$$
<sup>32</sup>

**Definition.** Given a performance index  $Q$ , say that  $u$  at  $w$  is *locally at least as inclined to invest in  $Q$ -performers* as  $v$  at  $w'$  if there exists  $\bar{q}$ , such that

---

<sup>32</sup>The requirement that  $r \in \mathcal{R}_1$  will be important in this setting due to the assumption of homogeneity of degree 0, since for any  $r$  with expected positive net returns,  $q > 0$ , and any agent, there exists a small enough  $\lambda > 0$  such that  $(q, \lambda \cdot r)$  will be accepted.

for all for all  $\bar{q} > q > 0$  and  $\delta > 0$  there exists  $\epsilon > 0$  with

$$\begin{aligned} 0 &\leq \sup_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is rejected by } u \text{ at } w\} \\ &\leq \inf_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is accepted by } v \text{ at } w'\} + \delta. \end{aligned}$$

The interpretation is as follows: for transactions with expected net return of  $q > 0$ , if  $v$  at  $w'$  is willing to invest in some local transaction, then  $u$  at  $w$  is willing to invest in any local transaction that performs significantly (by  $\delta$ ) better according to  $Q$ .

Next, I require that the relation locally at least as inclined to invest in  $Q$ -performers is reflexive.

**Axiom** (Reflexivity). For all  $u$  and  $w$ ,  $u$  at  $w$  is locally at least as inclined to invest in  $Q$ -performers as  $u$  at  $w$ .

**Definition.**  $u$  is globally inclined to invest in  $Q$ -performers at least as  $v$  if for all  $w, w'$ ,  $u$  is locally inclined to invest in  $Q$ -performers at wealth  $w$  at least as  $v$  at wealth  $w'$ .

**Axiom** (Strong global consistency). For every  $w \in \mathbb{R}$ ,  $q > 0$ , for every  $u$  and  $v$ , and every  $r, r' \in \mathcal{R}_1$ , if  $u$  is inclined to invest in  $Q$ -performers at least as  $v$ ,  $v$  accepts  $(q, r)$  at  $w$ , and  $Q(r') > Q(r)$ , then  $u$  accepts  $(q, r')$  at  $w$ .

The axiom roughly says that if an agent that cares less about  $Q$ -performance is willing to invests  $q$  in a market, it must be the case that an agent who cares more about  $Q$ -performance would be willing to invest the same amount when the market performs better.

### 1.6.2 Results

**Definition.** The *generalized Sharpe ratio* is defined as

$$P_{r_f}^{AS}(r) := P^{AS}(r - r_f) = \frac{\mathbb{E}[r - r_f]}{Q^{AS}(r - r_f)}.$$

*Continuity.* An index  $Q$  is continuous if for all  $r_f > -1$ ,  $Q_{r_f}(r_n) \rightarrow Q_{r_f}(r)$  whenever  $\{r_n\}$  and  $r$  are uniformly bounded market returns, and  $\{r_n\}$  converges to  $r$  in probability.

**Theorem 1.8.**  *$Q$  is a continuous index of performance that satisfies global consistency, reflexivity, translation invariance, monotonicity and homogeneity iff it is a continuous increasing transformation of  $P^{AS}(\cdot)$ .*

*Proof.* See appendix. □

*Remark.* On the domain of normally distributed market returns,  $P^{AS}$  is ordinally equivalent to the Sharpe ratio.

*Remark.*  $P^{AS}$  is increasing in odd distribution moments, and decreasing in even distribution moments.

**Proposition 1.3.**  *$P^{AS}$  is monotonic with respect to stochastic dominance in the presence of risk-free asset.*

*Proof.* If  $r_1$  dominates  $r_2$  in the presence of  $r_f$ , then there exist  $\alpha, \beta > 0$  such that  $\alpha r_1 + (1 - \alpha)r_f$  stochastically dominates  $\beta r_2 + (1 - \beta)r_f$ . There is no loss of generality in assuming that  $r_f = 0$  and  $\mathbb{E}[r_1] = \mathbb{E}[r_2]$ . With this assumption, the above implies  $\alpha r_1$  stochastically dominate  $\beta r_2$ . The monotonicity of  $Q^{AS}$  thus implies that  $Q^{AS}(\alpha r_1) < Q^{AS}(\beta r_2)$ , and stochastic

dominance implies  $\mathbb{E}[\alpha r_1] \geq \mathbb{E}[\beta r_2]$ . Altogether, these results imply

$$P_0^{AS}(r_1) = \frac{\mathbb{E}[r_1]}{Q^{AS}(r_1)} = \frac{\mathbb{E}[\alpha r_1]}{Q^{AS}(\alpha r_1)} > \frac{\mathbb{E}[\beta r_2]}{Q^{AS}(\beta r_2)} = \frac{\mathbb{E}[r_2]}{Q^{AS}(r_2)} = P_0^{AS}(r_2)$$

as required.  $\square$

**Corollary 1.7.**  *$Q$  is a continuous index of performance that satisfies global consistency, reflexivity, translation invariance, and monotonicity with respect to stochastic dominance in the presence of risk-free asset iff it is a continuous increasing transformation of  $P^{AS}(\cdot)$ .*

*Proof.* Follows from Claim 1.5 and Theorem 1.8.  $\square$

### 1.6.3 The Demand for Market Transactions

The next proposition provides a partition of market transactions into three: “attractive,” “unattractive” and ones about which the decision depends on wealth effects.

**Proposition 1.4.** *If  $\frac{q}{P^{AS}(g)} > \sup_w \{\rho_u^{-1}(w)\}$ , then  $u$  rejects  $(q, g)$  at any wealth level. If  $\frac{q}{P^{AS}(g)} < \inf_w \{\rho_u^{-1}(w)\}$ , then  $u$  accepts  $g$  at any wealth level.*

Next, I show that diversification makes transactions more desirable and that a property analogous to compound gambles holds.

**Proposition 1.5.** *Fix  $r_f$ . If  $g, h \in \mathcal{R}_{r_f+1}^{r_f}$  are such that  $(q, g)$  and  $(q, h)$  are accepted by  $u$  at any wealth level, then  $u$  accepts  $(q, \alpha g + (1 - \alpha)h)$  for all  $\alpha \in (0, 1)$  at any wealth level.*

*Proof.* From proposition 1.4 as  $P^{AS}(\alpha g + (1 - \alpha)h) \geq \min\{P^{AS}(g), P^{AS}(h)\}$ , by the properties of  $Q^{AS}$ .  $\square$



**Proposition 1.6.** Fix  $r_f$ , and let  $g, h \in \mathcal{R}_{r_f+1}^{r_f}$  be such that  $(q, g)$  and  $(q, h)$  are accepted by  $u$  at any wealth level, then if  $g$  and  $h$  are independent then  $u$  accepts  $(2q, \frac{1}{2}g + \frac{1}{2}h)$  at any wealth level.

*Proof.* From proposition 1.4 as  $P^{AS}(\frac{1}{2}g + \frac{1}{2}h) \geq 2 \cdot \min\{P^{AS}(g), P^{AS}(h)\}$ , by the properties of  $Q^{AS}$ .  $\square$

This proposition implies the analogue to Samuelson's theorem for the case where a risk-free asset exists.

**Example.** (*The demand for market portfolios*). Cabrales *et al.* (2014) use the estimates of risk aversion from Dohmen *et al.* (2011) to deduce that for relevant wealth levels a large fraction of the developed world population (importantly, not the very poor or the very rich) could be characterized by  $1.8 \cdot 10^{-6} < \rho_u < 5 \cdot 10^{-4}$ . Kadan and Liu (2014) use historical monthly return data from the American market and estimate  $E[r - r_f]$  by .406 and  $\frac{1}{R^{AS}}$  by .038 suggesting an estimated value of  $\frac{.406}{.038} \approx 10.69$  for  $P^{AS}$ . Based on these estimates, a policy maker may inform individuals that if they do not invest in the market they will (probably) be better-off by purchasing a well diversified portfolio with expected return of  $q$  where  $\frac{q}{10.69} < (5 \cdot 10^{-4})^{-1} = 2000$ , or, approximately,  $q < 20000$ . Finally, using the estimate for expected net return, this bound suggests that an exposure of less than  $\frac{20,000}{.406} \approx \$50000$  to a well diversified portfolio of American shares is better than holding just risk-free assets. An upper bound can also be suggested: investing more than  $\frac{(1.8 \cdot 10^{-6})^{-1} \cdot 10.69}{.406} \approx \$13.8$  million is dominated by opting out of the market.<sup>33</sup>

---

<sup>33</sup>For the upper bound I make the standard assumption that utilities present (weakly) decreasing absolute risk aversion.

**Example.** In the same setting, consider a policy maker who considers levying a tax on risky investment. Using the above estimates for risk aversion, and recalculating  $P^{AS}$  for the after tax return, the policy maker can derive an upper bound over possible tax revenues.

## 1.7 A Consistent Index of Relative Riskiness

This section presents an application for the setting of multiplicative gambles.

Define  $\mathcal{U} := \{u : \mathbb{R}_+ \rightarrow \mathbb{R} \mid \varrho_u(w) > 1 \ \forall w > 0\}$ , the set of (twice continuously differentiable) utility functions with relative risk aversion higher than that of the logarithmic utility function. Additionally, let  $\mathcal{H} := \{g \in \mathcal{G} \mid Q^{FH}(g) < 1\}$  be the set of gambles with FH riskiness smaller than 1. The following is a result of FH:

**Fact 1.1.**  $Q^{FH}(g) < 1 \iff \prod_i (1 + g_i)^{p_i} > 1 \iff \mathbb{E}[\log(1 + g)] > 0$ .

In what follows I will consider multiplicative gambles, so that now  $u$  *accepts*  $g$  at  $w$  if  $u(w + gw) > u(w)$ , and *rejects*  $g$  otherwise.<sup>34</sup> The interpretation of  $Q^{FH}(g) < 1$  is that gambles of the form  $wg$  are accepted by a logarithmic utility function at wealth  $w$ . Repeatedly accepting independent gambles with  $Q^{FH}(g) > 1$  would lead to bankruptcy with probability 1.

Adjusting the previous axioms to the current setting yields the following axioms for an index of (relative) riskiness  $Q : \mathcal{H} \rightarrow \mathbb{R}_+$ :

**Axiom** (Scaling).  $\forall \alpha > 0 \ \forall g \in \mathcal{H}, \ Q((1 + g)^\alpha - 1) = \alpha \cdot Q(g)$ .<sup>35</sup>

<sup>34</sup> $g$  can be interpreted as the return on some risky asset.

<sup>35</sup>Importantly, note that for every  $\alpha > 0$  if  $g \in \mathcal{H}$  then  $(1 + g)^\alpha - 1 \in \mathcal{H}$  by fact 1.1.

Similar to the homogeneity axiom, the scaling axiom embodies a cardinal interpretation.

**Definition** (Ordinally equivalent). Given an index of riskiness  $Q$ ,  $A_Q$  is *ordinally equivalent* to the coefficient of relative risk aversion  $\varrho$  if  $\forall u, v \in \mathcal{U} \ \forall w, w' > 0, \ A_Q(u, w) > A_Q(v, w') \iff \varrho_u(w) > \varrho_v(w')$ .<sup>36</sup>

**Theorem 1.9.** *If local consistency and scaling hold, then  $A_Q$  is ordinally equivalent to  $\varrho$ , and the relation at least as averse to  $Q$ -riskiness induces the same order as  $\varrho$ .*

*Proof.* omitted. □

**Axiom** (Global consistency). *For every  $u$  and  $v$  in  $\mathcal{U}$ , for every  $w > 0$  and every  $g$  and  $h$  in  $\mathcal{H}$ , if  $u \succ_Q v$ ,  $u$  accepts  $g$  at  $w$ , and  $Q(g) > Q(h)$ , then  $v$  accepts  $h$  at  $w$ .*

**Lemma 1.4.** *For any  $g \in \mathcal{H}$  there is a unique positive number  $S(g)$  such that  $\mathbb{E} \left[ (1 + g)^{-\frac{1}{S(g)}} \right] = 1$ .*

*Proof.* See appendix. □

**Definition.** The *index of relative riskiness*  $S$  of gamble  $g \in \mathcal{H}$  is implicitly defined by the equation  $\mathbb{E} \left[ (1 + g)^{-\frac{1}{S(g)}} \right] = 1$ .

**Theorem 1.10.**  *$S$  is the unique index of riskiness that satisfies local consistency, global consistency and scaling, up to a multiplication by a positive number.*

---

<sup>36</sup>Whenever the adaptation of a definition from the previous sections is clear, I omit it for brevity.

*Proof.* See appendix. □

As before, scaling is not always a desirable property. In what follows I omit this requirement.

**Axiom** (Strong global consistency). *For every  $u$  and  $v$  in  $\mathcal{U}$ , for every  $w > 0$  and every  $g$  and  $h$  in  $\mathcal{H}$ , if  $u \succeq_Q v$ ,  $u$  accepts  $g$  at  $w$ , and  $Q(g) > Q(h)$ , then  $v$  accepts  $h$  at  $w$ .*

**Theorem 1.11.** *If  $Q$  is a continuous index of relative-riskiness that satisfies monotonicity with respect to first order stochastic dominance and strong global consistency, and the relation at least as averse to  $Q$ -riskiness is reflexive, then  $Q$  is ordinally equivalent to  $S$ .*

*Proof.* See appendix. □

**Corollary 1.8.** *If  $Q$  is a continuous index of relative-riskiness that satisfies monotonicity with respect to first order stochastic dominance and strong global consistency, and the relation at least as averse to  $Q$ -riskiness is reflexive, then  $Q$  satisfies local consistency and  $A_Q$  is ordinally equivalent to  $\varrho$ .*

*Remark.* The monotonicity and continuity requirements could be replaced by other conditions as in Claim 1.4.

**Proposition 1.7.** *A gamble  $g$  with  $S(g) = c$  is rejected by  $u$  at  $w$  only if there exist some  $w'$  such that small gambles with  $S$ -riskiness of  $c$  are rejected. A gamble  $g$  with  $S(g) = c$  is accepted by  $u$  at  $w$  only if there exist some  $w'$  such that small gambles with  $S$ -riskiness of  $c$  are accepted.*

*Proof.* omitted. □

**Corollary 1.9.** *If  $S(g) > \sup_{w>0} \{A_S^{-1}(u, w)\} = \sup_{w>0} \{\varrho_u^{-1}(w)\}$  then  $u$  rejects  $g$  at any wealth level. If  $S(g) < \inf_{w>0} \{A_S^{-1}(u, w)\} = \inf_{w>0} \{\varrho_u^{-1}(w)\}$  then  $u$  accepts  $g$  at any wealth level.*

*Proof.* omitted. □

**Definition** (Compound gamble property). An index  $Q$  has the *compound gamble* property if for every compound gamble of the form  $f = (1 + g)(1 + \mathbf{1}_A h) - 1$ , where  $\mathbf{1}$  is an indicator,  $A$  is an event such that  $g$  is constant on  $A$  ( $g|_A \equiv x$  for some  $x$ ) and  $h$  is independent of  $A$ ,  $\max\{Q(g), Q(h)\} \geq Q(f) \geq \min\{Q(g), Q(h)\}$ .

**Proposition 1.8.**  *$S$  satisfies the compound gamble property. Thus, if  $g, h \in \mathcal{H}$  are independent, and  $\min\{S(g), S(h)\} > \sup_w \{\varrho_u^{-1}(w)\}$ , then a compound gamble of  $g$  and  $h$  will also satisfy the inequality. Additionally, if  $g, h \in \mathcal{H}$  are independent, and  $\max\{S(g), S(h)\} < \inf_w \{\varrho_u^{-1}(w)\}$ , then a compound gamble of  $g$  and  $h$  will also satisfy the inequality.*

*Proof.* omitted. □

**Axiom** (Generalized Samuelson property).  $\forall u, w > 0$   $S_Q^\infty(u, w) \geq \inf_{w>0} S_Q(u, w)$   
and  $R_Q^\infty(u, w) \leq \sup_{w>0} R_Q(u, w)$

**Theorem 1.12.**  *$S$  is the unique index of riskiness that satisfies the generalized Samuelson property, local consistency and scaling, up to a multiplication by a positive number.*

*Proof.* omitted. □

**Theorem 1.13.** *If  $Q$  satisfies the generalized Samuelson property, reflexivity, monotonicity with respect to first order stochastic dominance and continuity then  $Q$  is ordinally equivalent to  $S$ .*

*Proof.* omitted. □

## 1.8 Consistent Index of Delay

Similar to gambles, comparing cashflows which pay (require) different sums of money over several points in time is not a simple undertaking. Some pairs of cashflows may be compared using the partial order of *time-dominance* (Bøhren and Hansen, 1980; Ekern, 1981), which is the analogue of stochastic dominance in this setting. A cashflow  $c$  is *first-order time dominated* by  $c'$  if at any point in time the sum of money generated by  $c$  up to this point is lower than the sum that was generated by  $c'$ .<sup>37</sup> Bøhren and Hansen (1980) show that if  $c$  is first-order time dominated by  $c'$  then every agent with positive time preferences prefers  $c'$  to  $c$ . Positive time preferences mean that the agent prefers a dollar at time  $s$  to a dollar at time  $s + \Delta$  for all  $\Delta > 0$ . They also show that if  $c$  is *second-order time dominated* by  $c'$  then every agent with a decreasing and convex discounting function prefers  $c'$  to  $c$ .<sup>38</sup>

Time dominance is, however, a partial order. In this section, I use the consistency-motivated approach to derive a novel index for the delay embedded in an investment cashflow. The index I derive is new to the

---

<sup>37</sup>The sum may be negative, representing a required investment.

<sup>38</sup>As the definition of second-order time domination requires some notation, I choose to omit it, noting that it is analogous to second order stochastic dominance from the risk setting.

literature but it is related to the well-known internal rate of return. The index possesses several desirable properties similar to those of the AS index of riskiness. In particular, it is monotone with respect to time dominance.

### 1.8.1 Preliminaries

An *investment cashflow* is a sequence of outflows (investment) followed by inflows (return), and a sequence of times when they are conducted. Denote by  $c = (x_n, t_n)_{n=1}^N$  such a cashflow.<sup>39</sup> When  $x_n$  is positive the cashflow pays out  $x_n$  at time  $t_n$ , and when it is negative, an investment of  $|x_n|$  is required at  $t_n$ . Assume, without loss of generality, that  $t_1 < t_2 < \dots < t_N$ . Further, assume that  $x_1 < 0$  and  $\sum x_n > 0$ , so that some investment is required, and the (undiscounted) return is greater than the investment. This property implies that an agent that does not discount the future will accept any investment cashflow, while a sufficiently impatient agent will reject it. Let  $\mathcal{C}$  denote the collection of such cashflows, and  $\mathcal{C}_{t,\epsilon}$  be the collection of cashflows with  $t_1 \leq t \leq t_N$ , and  $t_N - t_1 < \epsilon$ .

An *index of delay* is a function  $T : \mathcal{C} \rightarrow \mathbb{R}_+$  from the collection of cashflows to the positive reals. A cashflow  $c$  is said to be more *T-delayed* than  $c'$  if  $T(c) > T(c')$ .

I consider a capital budgeting setting in which agent  $i$  discounts using a smooth schedule of positive instantaneous discount rates,  $r_i(t)$ .<sup>40,41</sup> Similar to  $\rho$  in the risk setting,  $r$  induces a complete order on all agent and time-point

---

<sup>39</sup>To keep notation simple, I avoid making the dependence of  $N$  on  $c$  explicit.

<sup>40</sup>An alternative interpretation may be a social planner with such time preferences (Foster and Mitra, 2003).

<sup>41</sup>For a discussion of this condition, see Bøhren and Hansen (1980) and references provided there.

pairs.<sup>42</sup> The *net present value (NPV)* of an investment cashflow  $c = (x_n, t_n)_{n=1}^N$  for the agent  $i$  at time  $t$  is

$$NPV(c, i, t) := \sum_n e^{-\int_t^{t_n} r_i(s) ds} x_n.$$

If  $NPV(c, i, t) > 0$  for some  $t$ , this inequality holds for any  $t$ . Agent  $i$  *accepts* cashflow  $c$  (at time  $t$ ) if  $NPV(c, i, t) > 0$  and *rejects* it otherwise.  $c$  could be thought of as a suggested shift to a baseline cashflow.

The following two definitions are crucial for applying the consistency motivated approach from the previous sections in order to present axioms for an index of delay. Given an index of delay  $T$ , an agent  $i$ , a time  $t$ , and  $\epsilon > 0$ :

**Definition.**  $R_T^\epsilon(i, t) := \sup \{T(c) \mid c \in \mathcal{C}_{t, \epsilon} \text{ and } c \text{ is accepted by } i\}$

**Definition.**  $S_T^\epsilon(i, t) := \inf \{T(c) \mid c \in \mathcal{C}_{t, \epsilon} \text{ and } c \text{ is rejected by } i\}$

$R_T^\epsilon(i, t)$  is the  $T$ -delay of the *most delayed* cashflow according to  $T$  that  $i$  is willing to accept, restricting the support of the cashflows to an  $\epsilon$ -ball around  $t$ .  $S_T^\epsilon(i, t)$  is the  $T$ -delay of the *least delayed* cashflow according to  $T$  which  $i$  rejects, again restricting the support of the cashflows to an  $\epsilon$ -ball around  $t$ .

**Definition.**  $i$  at  $t$  is *at least as averse to  $T$ -delay* as  $j$  at  $t'$  if for every  $\delta > 0$  there exists  $\epsilon > 0$  such that  $S_Q^\epsilon(j, t') \geq R_Q^\epsilon(i, t) - \delta$ .

The interpretation of  $i$  at  $t$  being at least as averse to  $T$ -delay as  $j$  at  $t'$  is that, at least for cashflows with a short horizon, if  $i$  accepts *any* short-horizon cashflow concentrated around  $t$  with a certain level of  $T$ -delay,  $j$  accepts *all*

---

<sup>42</sup>Importantly,  $r$  is not a common *interest rates* path as in Debreu (1972).



short-horizon cashflows which are significantly (by at least  $\delta$ ) less delayed according to  $T$  and are concentrated around  $t'$ . Alternatively, if  $j$  rejects *any* short-horizon cashflow that is concentrated around  $t'$  and has a certain level of  $T$ -delay,  $i$  rejects *all* short horizon cashflows which are significantly (by at least  $\delta$ ) more  $T$ -delayed and are concentrated around  $t$ .

The following definitions will also prove useful:

**Definition.**  $R_T(i, t) := \lim_{\epsilon \rightarrow 0^+} R_T^\epsilon(i, t)$

**Definition.**  $S_T(i, t) := \lim_{\epsilon \rightarrow 0^+} S_T^\epsilon(i, t)$

Roughly speaking,  $R_T(i, t)$  is the  $T$ -delay of the most  $T$ -delayed short-horizon cashflow that is concentrated around  $t$  and accepted by  $i$ , and  $S_T(i, t)$  is the  $T$ -delay of the least  $T$ -delayed short-horizon cashflow that is concentrated around  $t$  and rejected by  $i$  at  $t$ . As before, the *coefficient of local aversion to  $T$ -delay of  $i$  at  $t$*  is therefore defined as

$$A_T(i, t) := \frac{1}{R_T(i, t)},$$

noting that all of the results would hold for  $\frac{1}{S_T(i, t)}$  as well.

### 1.8.2 The Index

The following axioms are an adaptation of the axioms used in Theorem 1.3 for the current setting. They are used for presenting the analogue of this theorem, as well as the analogue of Theorem 1.2. Theorem 1.14 provides conditions under which there is only one order of local aversion to delay and it corresponds to the instantaneous discount rate.

**Axiom** (Translation invariance).  $T\left((x_n, t_n + \lambda)\right)_{n=1}^N = T\left((x_n, t_n)\right)_{n=1}^N$  for any cashflow and any  $\lambda > 0$ .

Translation invariance of  $T$  means that  $T$ -delay is a time expression, like “in a week” or “a year before,” and it does not depend on the start date. In contrast, the interpretation of expressions such as “this Tuesday” depends critically on whether they are said on Friday or Monday. This will be the only “new” requirement in the current setting; all other axioms are adaptations of the axioms from the risk settings to the current one.

**Axiom** (Homogeneity (of degree  $k$  in dates)). For any cashflow with  $t_1 = 0$ , for any  $\lambda > 0$ ,  $T\left((x_n, \lambda \cdot t_n)\right)_{n=1}^N = \lambda^k \cdot T\left((x_n, t_n)\right)_{n=1}^N$  for some  $k > 0$ .

Homogeneity of degree 1 in dates, when combined with translation invariance, represents the notion that if each payment in the cashflow is conducted twice as late relative to the first period of investment, then the entire cashflow is twice as delayed relative to that time. This is a strong cardinal assumption and I later discuss its removal.

**Axiom** (Local consistency).  $\forall i \ \forall t \ \exists \lambda > 0 \ \forall \delta > 0 \ \exists \epsilon > 0 \ R_T^\epsilon(i, t) - \delta < \lambda < S_T^\epsilon(i, t) + \delta$ .

Local consistency says that cashflows which are “local” with respect to  $t$  that are significantly less  $T$ -delayed than some cut-off level are always accepted by  $i$ , and that ones significantly more  $T$ -delayed than the cutoff are always accepted. Lemma A.13 in the appendix shows that whenever homogeneity is satisfied, local consistency implies that  $0 < S_T(i, t) = R_T(i, t) < \infty$ . This means, that for “local” cashflows  $T$  is sufficient information to determine

an agent's optimal behavior. In other words, the decisions of agents are consistent with the index, on small domains.

**Definition** (Reflexivity). The relation at least as averse to  $T$ -delay is *reflexive* if for all  $i$  and  $t$ ,  $i$  at  $t$  is at least as averse to  $T$ -delay as  $i$  at  $t$ .

**Proposition 1.9.** *If  $T$  satisfies local consistency, then the relation locally at least as averse to  $T$ -delay is reflexive.*

**Definition** (Ordinally equivalent). Given an index of delay  $T$ ,  $A_T$  is *ordinally equivalent* to the instantaneous discount rate  $r$  if  $\forall i, j, \forall t, t' A_T(i, t) > A_T(j, t') \iff r_i(t) > r_j(t')$ .

**Theorem 1.14.** *If  $T$  satisfies local consistency, homogeneity and translation invariance, then  $A_T$  is ordinally equivalent to  $r$ , and the relation at least as averse to  $T$ -delay induces the same order as  $r$ .*

*Proof.* See appendix. □

*Remark.* All axioms in Theorem 1.14 are essential: As the following examples demonstrate, omitting any admits indices to which the coefficient of local aversion is not ordinally equivalent to  $r$ , and the relation at least as averse to  $T$ -delay does not induce the same order as  $r$ .

**Example 1.12.**  $T \equiv 5$  satisfies local consistency and translation invariance, but it does not satisfy homogeneity of degree  $k > 0$ . The local aversion to this index induces the trivial order and  $A_T \equiv \frac{1}{5}$ .

**Example 1.13.**  $T := t_2 - t_1$  satisfies homogeneity and translation invariance, as  $\lambda t_2 - \lambda \cdot 0 = \lambda(t_2 - 0)$  and  $t_2 - t_1 = (t_2 + \lambda) - (t_1 + \lambda)$ . Local consistency is,

however, violated. The local aversion to this index induces the trivial order and  $A_T \equiv \infty$ .

Example 1.14 will demonstrate that without translation invariance the inference is not necessarily correct. The following two definitions prove useful for the example as well as for the statement and proof of Theorem 1.16.

**Definition.** The *internal rate of return* (IRR) of an investment cashflow  $c = (x_n, t_n)_{n=1}^N$ , written  $\alpha(c)$ , is the unique positive solution to the equation  $\sum_n e^{-\alpha t_n} x_n = 0$ .

Existence and uniqueness follow from Lemma A.12 which generalizes the result of Norström (1972) who had shown that investment cashflows have a unique positive IRR in the discrete setting. For general cashflows, multiple solutions to the equation defining the internal rate of return may exit.<sup>43</sup>

**Definition.** For a cashflow  $c$ ,  $D(c) := \frac{1}{\alpha(c)}$  is the inverse of the IRR of the cashflow.

**Example 1.14.** Consider the index of delay

$$T(c) = \begin{cases} D(c) & \text{if } t_1 < 3 \text{ or } 5 < t_N \\ (t_1 - 2) \cdot D(c) & \text{if } 3 \leq t_1 \leq 4 \\ (6 - t_1) \cdot D(c) & \text{if } 4 \leq t_1 \leq 5. \end{cases}$$

It is homogeneous since it coincides with  $D$  on the relevant domain. It is locally consistent since  $D$  is, a fact which will be proved later, and since for

---

<sup>43</sup>In addition, phenomena with the flavor of reswitching might arise (Levhari and Samuelson, 1966), as discussed in Footnote 45.

any  $t$ , in small environments of  $t$  the index is approximately equal to  $C \cdot D(\cdot)$  for some  $C = C(t)$ . Now, consider an agent,  $i$ , with a constant discount rate  $r_i(t) \equiv r$ . For  $t = 4$ , the coefficient of  $T$ -delay aversion of the agent is not equal to the coefficient of  $T$ -delay aversion for the same agent at  $t = 1$ . But  $r_i(\cdot)$  is constant by construction. It is also the case that  $i$  at  $t = 4$  is not at least as averse to  $T$ -delay as  $i$  at  $t = 1$ .

**Theorem 1.15.** (i) *There exists a continuum of translation invariant, locally consistent, homogeneous of degree 1 indices of delay to which the local aversion equals to  $r$ .* (ii) *Moreover, some of these indices are not monotone with respect to first order time dominance.*<sup>44</sup>

*Proof.* See appendix. □

**Definition** (Globally more  $T$ -delay averse).  $i$  is *Globally at least as  $T$ -delay averse* as  $j$  (denoted  $j \lesssim_T i$ ) if for every  $t$  and  $t'$ ,  $i$  at  $t$  is at least as averse to  $T$ -delay as  $j$  at  $t'$ .  $i$  is *globally more  $T$ -delay averse* than  $j$  (denoted by  $j <_T i$ ) if  $j \lesssim_T i$  and not  $i \lesssim_T j$ .

This definition generates a partial order over agents, based on their preferences and on the index of delay. As before, global consistency is an important part of the approach.

**Axiom** (Global consistency). *If  $j <_T i$ ,  $T(c) < T(c')$ , and  $i$  accepts  $c'$ , then  $j$  accepts  $c$ .*<sup>45</sup>

---

<sup>44</sup> $T$  satisfies monotonicity with respect to first order time dominance if  $T(c) < T(c')$  whenever  $c$  time dominates  $c'$ .

<sup>45</sup>The use of acceptance and rejection allows me to avoid the *reswitching* problem of the famous *Cambridge capital controversy* (See Cohen and Harcourt (2003) for an extensive review). In contrast to choices between two cashflows, which, in general, may not be

**Theorem 1.16.**  $D^k(\cdot)$  is the unique index of delay that satisfies local consistency, global consistency, homogeneity of degree  $k > 0$  and translation invariance, up to a multiplication by a positive number.

*Proof.* See appendix. □

The homogeneity axiom is not necessarily appealing in the current setting. In what follows, it will be removed and replaced with less demanding conditions: monotonicity with respect to first order time dominance and continuity. As in previous sections, Example 1.15 below shows that these conditions are not enough to pin down desirable indices. Hence, I will require a slightly stronger version of global consistency but, as before, will replace the local consistency requirement with the weaker requirement of reflexivity.

**Definition** (Continuity). An index of delay is continuous if  $T(c_n) \rightarrow T(c)$  whenever  $\{c\} \cup \{c_n\} \subset \mathcal{C}$ , random variables with distribution  $\left(\frac{|x_i^n|}{\sum_{i|x_i>0} |x_i^n|}, t_i^n\right)$  and  $\left(\frac{|x_i^n|}{\sum_{i|x_i\leq 0} |x_i^n|}, t_i^n\right)$  converge in probability to  $\left(\frac{|x_i|}{\sum_{i|x_i>0} |x_i|}, t_i\right)$  and  $\left(\frac{|x_i|}{\sum_{i|x_i\leq 0} |x_i|}, t_i\right)$  respectively if all random variables are uniformly bounded and  $\sum x_i^n$  converges to  $\sum x_i$ .

**Example 1.15.** Consider the index

$$T(c) := 1 + \sum_{j|x_j>0} \frac{|x_j|t_j}{\sum_{i|x_i>0} |x_i|} - \sum_{j|x_j\leq 0} \frac{|x_j|t_j}{\sum_{i|x_i\leq 0} |x_i|}.$$

It is well-defined and positive as the first summation is a weighted average of greater numbers and both summations are non-degenerate, by the definition

---

monotonic in the discount rate, acceptance and rejection decisions of investment cashflows are monotonic in these rates. This is shown in Lemma A.12 in the appendix.

of investment cashflow. It is translation invariant since adding  $t$  to all  $t_i$ 's increases both summations by  $t$ . Continuity follows directly from the definition. Homogeneity of degree 0 in payoffs holds as well, since weights are not changed when all  $x_i$ 's are multiplied by a positive number. Local consistency holds since both summations converge to  $t$ , when considering smaller and smaller environments of  $t$ , and so  $R_T \equiv S_T \equiv 1$ . Hence, the coefficient of local aversion to  $T$ -delay is identically equal to 1, and every  $i$  is globally at least as averse to  $T$ -delay as any  $j$ . Thus, the relation more averse to  $T$ -delay is empty and global consistency is automatically satisfied.

**Axiom** (Strong global consistency). *If  $j \lesssim_T i$ ,  $T(c) < T(c')$ , and  $i$  accepts  $c'$ , then  $j$  accepts  $c$ .*

**Theorem 1.17.** *If  $T$  is a continuous index of delay that satisfies monotonicity with respect to first order time dominance, translation invariance and strong global consistency, and the relation at least as averse to  $T$ -delay is reflexive, then  $T$  is ordinally equivalent to  $D$ .*

*Proof.* See appendix. □

**Corollary 1.10.** *If  $T$  is a continuous index of delay that satisfies monotonicity with respect to first order time dominance, translation invariance and strong global consistency, then  $T$  is locally consistent and  $A_T$  is ordinally equivalent to  $r$ .*

*Remark.* The monotonicity requirement in the theorem could be replaced by each of the following conditions:

- (a) Satisfying the ordinal content of homogeneity

(b) Monotonicity with respect to delaying the first investment period, leaving the rest of the periods unchanged

In such case, monotonicity with respect to first order time dominance will be a result, not an assumption.<sup>46</sup>

*Remark.* The partial order globally at least as  $D$ -delay averse is refined by the partial orders of delay aversion of Horowitz (1992) and Benoît and Ok (2007).

### 1.8.3 $D$ -Delay Aversion and the Demand for Investment Cash-flows

**Proposition 1.10.** *A cashflow  $c = (x_n, t_n)_{n=1}^N$  with  $D(c) = b$  is rejected by  $i$  only if there exist some  $t \in [t_1, t_N]$  such that small cashflows with  $D$  of  $b$  are rejected. A cashflow  $c = (x_n, t_n)_{n=1}^N$  with  $D(c) = b$  is accepted by  $i$  only if there exist some  $t \in [t_1, t_N]$  such that small cashflows with  $D$  of  $b$  are accepted.*

*Proof.* See appendix. □

**Corollary 1.11.** *If  $D(c) > \sup_t \{A_D^{-1}(i, t)\} = \sup_t \{r_i^{-1}(t)\}$  then  $i$  rejects any translation of  $c$ . If  $D(c) < \inf_t \{A_D^{-1}(i, t)\} = \inf_t \{r_i^{-1}(t)\}$  then  $i$  accepts any translation of  $c$ .*

Similar to results in previous sections, the corollary suggests a partition of the class of cashflows into three: ones which the agent never accepts, ones which are always accepted, and ones whose acceptance or rejection may not be determined.

---

<sup>46</sup>The continuity assumption could also be relaxed.



**Definition** (Compound cashflow property). An index  $T$  has the *compound cashflow* property if for every compound cashflow of the form  $f = c + c'$ ,<sup>47</sup> where  $c, c'$  and  $f$  are investment cashflows  $\max\{T(c), T(c')\} \geq T(f) \geq \min\{T(c), T(c')\}$ .

**Proposition 1.11.**  $D$  satisfies the compound cashflow property. Thus, if  $c, c', c + c' \in \mathcal{C}$  and  $\min\{D(c), D(c')\} > \sup_t \{r_i^{-1}(t)\}$ , then  $c + c'$  also satisfies the inequality; and if  $c, c', c + c' \in \mathcal{C}$  and  $\max\{D(c), D(c')\} < \inf_t \{r_i^{-1}(t)\}$  then  $c + c'$  also satisfies the inequality.

*Proof.* See appendix. □

**Axiom** (Generalized Samuelson property).  $\forall i \ S_T^\infty(i) \geq \inf_t S_T(i, t)$  and  $R_T^\infty(i) \leq \sup_t R_T(i, t)$ .

**Theorem 1.18.**  $D^k(\cdot)$  is the unique index of delay that satisfies the generalized Samuelson property, local consistency, homogeneity of degree  $k > 0$  and translation invariance, up to a multiplication by a positive number.

*Proof.* Omitted. □

**Theorem 1.19.** If  $T$  satisfies the generalized Samuelson property, translation invariance, reflexivity, monotonicity with respect to first order time dominance and continuity then  $T$  is ordinally equivalent to  $D$ .

*Proof.* Omitted. □

---

<sup>47</sup>The interpretation of  $c + c'$  is that all of the payoffs which are dictated by each of the cashflows takes place at the times they dictate. If both require a payoff at the same time point, the payoffs are added up.

#### 1.8.4 Other Properties of $D$ and a Comparison with $Q^{AS}$

This section discusses some properties of the index of delay  $D$  and demonstrates the close connection it has with the AS index of riskiness. The IRR is a counterpart of *the rate of return over cost* suggested by Fisher (1930) as a criterion for project selection almost a century ago. Later, some economists dismissed this criterion, arguing that the NPV was superior in comparing pairs of cashflows. Yet, others mentioned that this criterion has the benefit of objectivity, in that it does not require the value judgment of setting the future discount rates (Turvey, 1963). For example, Stalin and Nixon would agree on the IRR of an investment even though they might disagree on its NPV.<sup>48</sup>

Just like the AS-riskiness of a gamble depends “*on its distribution only—and not on any other parameters, such as the utility function of the decision maker or his wealth*” (Aumann and Serrano, 2008),  $D$  depends solely on the cashflow, and not on any agent specific properties. In this sense,  $D$  is an objective measure of delay. In particular,  $D$  is independent of the date when the cashflow is considered. That is, the  $D$ -delay embedded in an investments cashflow is independent of the time when it is considered.

$D$  is homogeneous of degree 0 in payoffs and unit free. This means, for example, that the  $D$ -delay of two cashflows denominated in different currencies may be compared without knowledge of the exchange rate. This stands in contrast to the AS index of riskiness which is homogeneous of degree

---

<sup>48</sup>This resembles the point made by Hart (2011) that in general there are many pairs of agents and pairs of gambles such that each agent accepts a different gamble and rejects the other – our axioms only compare very specific pairs of agents.

1 in payoffs, but does not depend on timing. The property is analogous to the property of  $Q^{AS}$ , according to which “diluted” gambles inherit the riskiness of the original gamble. For  $p \in (0, 1)$  a  $p$ -*dilution* of the gamble  $g$  takes the value of the gamble with probability  $p$  and 0 with probability  $1 - p$ , independently of the gamble. The reason why this analogy is correct is that in the current setting, times are the parallel of payoffs from the risk setting, while payoffs are the parallel of probabilities, as demonstrated by the remark at the end of this section.

Another property that  $D$  and  $Q^{AS}$  share is monotonicity.  $Q^{AS}$  is monotonic with respect to first and second order stochastic dominance. The analogous property for cashflows is *time-dominance* (Bøhren and Hansen, 1980; Ekern, 1981). Proposition 3 of Bøhren and Hansen (1980) implies that  $D$  is monotonic with respect to time-dominance of any order.

There are other similarities between the measurement of delay and risk. *Value at Risk* (VaR) is a family of indices commonly used in the financial industry (Aumann and Serrano, 2008). VaR indices depend on a parameter called the *confidence level*. For example, the VaR of a gamble at the 95 percent confidence level is the largest loss that occurs with probability greater than 5 percent. Unlike the AS index, VaR is unaffected by tail events or rare-disasters, extremely negative outcomes that occur with low probability. In the context of project selection, *Turvey (1963)* mentions that “*the Pay-off Period, the number of years which it will take until the undiscounted sum of the gains realized from the investment equals its capital cost,*” was used by practitioners in the West and in Russia. He adds that “[p]ractical men in industries with long-lived assets have perforce been made aware of the

*deficiencies of this criterion and have sought to bring in the time element.”*

The pay-off period criterion, unlike the index of delay, suffers from deficiencies similar to those of VaR. For example, shifting early or late payoffs does not change its value. In fact, recalling that times in the current setting are the parallel of payoffs in the risk setting, the lesson learned by the investors in long-lived assets should apply to investors in risky assets with distant tail events.

$Q^{AS}$  is much more sensitive to the loss side of gambles than it is to gains. Analogously  $D$  is more sensitive to early flows than it is to later ones. This follows from the properties of the exponential function in the definition of the IRR. Additionally, both  $D$  and  $Q^{AS}$  are continuous in their respective spaces.

Finally, to clarify the analogies I made between probabilities and payoffs, and between payoffs and times, I present a reinterpretation of the AS index of riskiness in terms of the delay embedded in a (non-investment) cashflow.

*Remark.* Given a gamble  $g := (g_j, p_j)$ , a cashflow which requires an investment of one dollar at  $t = 0$  and pays-out  $p_j$  at time  $g_j$  has a unique positive IRR whose inverse equal to  $Q^{AS}(g)$ .<sup>49</sup>

To see this, recall that for a cashflow  $c = (x_n, t_n)_{n=1}^N$  the (unique) positive IRR is the (unique) positive solution to the equation  $\sum_n e^{-\alpha t_n} x_n = 0$ , when it exists. Noting that at  $t = 0$ ,  $e^{-\alpha t} = 1$  and that the above cashflow requires an investment of one dollar at  $t = 0$ , the corresponding equation could be

---

<sup>49</sup>This is not the unique IRR as 0 is also a solution of the defining equation.

written as

$$-1 + \sum_n e^{-\alpha g_n} p_n = 0,$$

which could be expressed as

$$\mathbb{E}[e^{-\alpha g}] = 1.$$

But  $Q^{AS}(g)$  is the inverse of the unique positive  $\alpha$  which solves the equation.

For general cashflows, multiple solutions to the equation defining the internal rate of return may exist. Interestingly, both Arrow and Pratt took interest in finding simple conditions that would rule out this possibility (Arrow and Levhari, 1969; Pratt and Hammond, 1979). A corollary of the previous remark is that cashflows of the above form have a unique positive IRR.

## 1.9 A Consistent Index of the Appeal of Information Transactions

Similar to the previous settings, generating a sensible complete ranking of information structures is an illusive undertaking. In some settings, certain information may be vital, while in others it will not be very important. The implication is that it is not possible to rank all information structures so that higher ranked structures are preferred to lower ranked ones by all agents at every decision making problem. Some pairs of information structures may, however, be compared in this manner. Blackwell's (1953) seminal paper shows that one information structure is preferred to another by all agents in

all settings if and only if the latter is a garbling of the prior,<sup>50</sup> that is, if one is a noisy version of the other. But this order is partial and cannot be used to compare many pairs of information structures.

The difficulty in generating a complete ranking which is independent of agents' preferences is discussed by Willinger (1989) in his paper which studies the relation between risk aversion and the value of information. Willinger (1989) discusses his choice of using the *expected value of information* (EVI) or "*asking price*" which was defined by LaValle (1968). The EVI measures a certain decision maker's willingness to pay for certain information, and so, "... *the difficulty of defining a controversial continuous variable representing the 'amount of information' can be avoided.*"

Cabrales *et al.* (2013) tackle this difficulty using an approach in the spirit of Hart (2011). They restrict attention to a decision problem of information acquisition by investors in a model a la Arrow (1972) and define an order which they name *uniform investment dominance*, which turns out to be a complete order over all *information structures*. In a separate paper, these authors take an approach in the spirit of AS, and axiomatically derive a different index for the appeal of *information transactions* (Cabrales *et al.*, 2014). Both approaches lead to orders which refine the order suggested by Blackwell (1953), however, they depend on the (unique, fixed, common) prior of the decision makers which are considered.

In this section, I study a problem of information acquisition by investors using the same techniques as in previous sections. I show that the coefficient of local taste for  $Q$ -informativeness is equal to the inverse of ARA when  $Q$  is

---

<sup>50</sup>A simple proof is provided in Leshno and Spector (1992).

one of these two prominent indices, and that the unique index which satisfies local consistency, global consistency, and a homogeneity axiom is the index of appeal of information transactions (Cabrales *et al.*, 2014). As always, an ordinal version of this result, which does not assume homogeneity, is provided. The section ends with a discussion of the prior-free implications of the index of appeal of information transactions (which is prior-dependent).

### 1.9.1 Preliminaries

This section follows closely Cabrales *et al.* (2014). I consider agents with concave and twice continuously differentiable utility functions who have some initial wealth and face uncertainty about the state of nature. There are  $K \in \mathbb{N}$  states of nature,  $\{1, \dots, K\}$ ,<sup>51</sup> over which the agents have the prior  $p \in \Delta(K)$  which is assumed to have a full support.

The set of *investment opportunities*  $B^* = \left\{ b \in \mathbb{R}^K \mid \sum_{k \in K} p_k b_k \leq 0 \right\}$ , consists of all no arbitrage assets. In particular it includes the option of inaction. The reference to the members of  $B^*$  as no arbitrage investment opportunities attributes to  $p_k$  an additional interpretation as the price of an Arrow-Debreu security that pays 1 if the state  $k$  is realized and nothing otherwise. Hence,  $p$  plays a dual role in this setting. When an agent with initial wealth  $w$  chooses investment  $b \in B^*$  and state  $k$  is realized, his wealth becomes  $w + b_k$ .

Before choosing his investment, the agent has an opportunity to engage in an *information transaction*  $a = (\mu, \alpha)$ , where  $\mu > 0$  is the cost of the transactions, and  $\alpha$  is the information structure representing the information

---

<sup>51</sup>With a slight abuse of notation, I also denote  $\{1, \dots, K\}$  by  $K$ . The meaning of  $K$  should be clear from the context.

that  $a$  entails. To be more precise,  $\alpha$  is given by a finite set of signals  $S_\alpha$  and probability distributions  $\alpha_k \in \Delta(S_\alpha)$  for every  $k \in K$ . When the state of nature is  $k$ , the probability that the signal  $s$  is observed equals  $\alpha_k(s)$ . Thus, the information structure may be represented by a stochastic matrix  $M_\alpha$ , with  $K$  rows and  $|S_\alpha|$  columns, and the total probability of the signals is given by the vector  $p_\alpha := p \cdot M_\alpha$ . For simplicity, assume that  $p_\alpha(s) > 0$  for all  $s$ , so that each signal is observed with positive probability. Further, denote by  $q_k^s$  the probability the agent assigns to state  $k$  conditional on observing the signal  $s$ , using Bayes' law. Note that although my notation does not indicate it,  $(q_k^s)_{k=1}^K = q^s \in \Delta(K)$  depends on  $\alpha$  and the prior  $p$ .

The transaction  $a$  is said to be *excluding* if for every  $s$  there exists some  $k$  such that  $q_k^s = 0$ . This means that for every signal the agent receives, he knows that some states will not be realized (allowing him to generate arbitrarily large profits with certainty). Throughout, I will assume that information transactions are not excluding.

Agents are assumed to optimally choose an investment opportunity in  $B^*$  given their belief,  $q$ . Therefore, the expected utility of an agent with utility  $u$ , initial wealth  $w$  and beliefs  $q$  is

$$V(u, w, q) := \sup_{b \in B^*} \sum_k q_k u(w + b_k).$$

In case that the agent acquires no information, his beliefs are given by the prior  $p$ . Since the agent is risk averse, in such case his optimal choice is inaction. Hence,

$$V(u, w, p) = u(w).$$



Accordingly, an agent *accepts* an information transaction if

$$\sum_s p_\alpha(s) V(u, w - \mu, q^s) > V(u, w, p) = u(w)$$

and *rejects* it otherwise.

Denote by  $\mathcal{A}$  the class of information transactions described above. Additionally, denote by  $\mathcal{A}_\epsilon$  the sub-class of these information transactions such that  $\|p - q^s\|_\infty < \epsilon$  for all  $s$ . An *index of appeal* of information transactions is a function from the class of information transactions to the positive reals  $Q : \mathcal{A} \rightarrow \mathbb{R}_+$ . The index of appeal  $A$  suggested by Cabrales *et al.* (2014) is defined by

$$A(a) = -\frac{1}{\mu} \log \left( \sum_s p_\alpha(s) \exp(-d(p||q^s)) \right),$$

where

$$d(p||q) = \sum_k p_k \log \frac{p_k}{q_k}$$

is the *Kulback-Leibler divergence* (Kullback and Leibler, 1951).

Cabrales *et al.* (2013) suggest the *entropy reduction* as a measure of informativeness of an information structure for investors. It is defined by

$$I_e(\alpha) = H(p) - \sum_s p_\alpha(s) \cdot H(q^s),$$

where,

$$H(q) = -\sum_{k \in K} q_k \log(q_k).$$

In the current context, consider the index  $J_e$ , the *cost adjusted entropy reduction* defined by

$$J_e(\mu, \alpha) = \frac{I_e(\alpha)}{\mu}.$$

To apply the techniques from the previous sections, some more definitions are required. Given an index of informativeness  $Q$ , a utility function  $u$ , a wealth level  $w$  and  $\epsilon > 0$ :

**Definition.**  $R_Q^\epsilon(u, w) := \inf \{Q(a) \mid a \in \mathcal{A}_\epsilon \text{ and } a \text{ is accepted by } u \text{ at } w\}$

**Definition.**  $S_Q^\epsilon(u, w) := \sup \{Q(a) \mid a \in \mathcal{A}_\epsilon \text{ and } a \text{ is rejected by } u \text{ at } w\}$

$R_Q^\epsilon(u, w)$  is the  $Q$ -informativeness of the *least informative* accepted transaction according to  $Q$ , which is in  $\mathcal{A}_\epsilon$ .  $S_Q^\epsilon(u, w)$  is the  $Q$ -informativeness of the *most informative* rejected transaction according to  $Q$ , again restricting the support of the transactions to  $\mathcal{A}_\epsilon$ .

$u$  at  $w$  has *at least as much taste for  $Q$ -informativeness* as  $v$  at  $w'$  if for every  $\delta > 0$  there exists  $\epsilon > 0$  such that  $S_Q^\epsilon(u, w) \leq R_Q^\epsilon(v, w') + \delta$ .

The interpretation of  $u$  at  $w$  having at least as much taste for  $Q$ -informativeness as  $v$  at  $w'$  is that, at least for small transactions, if  $v$  at  $w'$  accepts *any* small transactions with a certain level of  $Q$ -informativeness,  $u$  at  $w$  accepts *all* small transactions which are significantly (by at least  $\delta$ ) more  $Q$ -informative. The following definitions will also prove useful:

**Definition.**  $R_Q(u, w) := \lim_{\epsilon \rightarrow 0^+} R_Q^\epsilon(u, w)$

**Definition.**  $S_Q(u, w) := \lim_{\epsilon \rightarrow 0^+} S_Q^\epsilon(u, w)$

$S_Q(u, w)$  is the  $Q$ -appeal of the most  $Q$ -appealing transaction that is rejected, and never provides a lot of information, in the sense that the posterior and the prior are close.<sup>52</sup> Finally, define the *coefficient of local*

---

<sup>52</sup>Note that in this setting the index is not independent of the prior  $p$ , even when the dependence is not made explicit by the notation I use.

taste for  $Q$ -informativeness of an agent  $u$  with wealth  $w$  as the inverse of  $S_Q(u, w)$ .

### 1.9.2 The Index

Theorem 1.20 is the analogue of Theorem 1.1 in the current context. It shows that the coefficient of local taste for  $Q$ -informativeness coincides with the inverse of  $\rho$  for the two indices of informativeness discussed above.<sup>53</sup>

**Theorem 1.20.** (i) For every  $u$  and  $w$ ,  $R_A(u, w) = S_A(u, w) = \rho_u(w)$ . (ii) For every  $u$  and  $w$ ,  $R_{J_e}(u, w) = S_{J_e}(u, w) = \rho_u(w)$ .

*Proof.* See appendix. □

**Corollary 1.12.** For  $Q \in \{A, J_e\}$   $u$  at  $w$  has at least as much taste for  $Q$ -informativeness as  $v$  at  $w'$  iff  $\rho_u(w) \leq \rho_v(w')$ .

The following two theorems are the analogues of Theorems 1.2 and 1.3.

**Axiom** (Homogeneity). There exists  $k > 0$  such that for every information transaction  $a = (\mu, \alpha)$  and every  $\lambda > 0$ ,  $Q(\lambda \cdot \mu, \alpha) = \frac{1}{\lambda^k} \cdot Q(a)$ .

The homogeneity axiom states that  $Q$  is homogeneous of degree  $-k$  in transaction prices. This axiom entails the cardinal content of the index. It is particularly interesting if  $k = 1$ . In this case, the units of the index could be interpreted as *information per dollar*.

**Axiom** (Local consistency).  $\forall u \forall w \exists \lambda > 0 \forall \delta > 0 \exists \epsilon > 0 R_Q^\epsilon(u, w) + \delta > \lambda > S_Q^\epsilon(u, w) - \delta$ .

---

<sup>53</sup>The relations between risk aversion and the taste for information have been discussed extensively in the literature (e.g. Willinger, 1989).

**Definition** (Reflexivity). The relation has at least as much taste for  $Q$ -informativeness is *reflexive* if for all  $u$  and  $w$ ,  $u$  at  $w$  has at least as much taste for  $Q$ -informativeness as  $u$  at  $w$ .

**Proposition 1.12.** *If  $Q$  satisfies local consistency, then the relation has at least as much taste for  $Q$ -informativeness is reflexive.*

**Theorem 1.21.** *Fix  $k > 0$ . If  $Q$  satisfies local consistency and homogeneity of degree  $-k$  in prices, then the coefficient of local taste for  $Q$ -informativeness is ordinally equivalent to  $\rho^{-1}$ , and the relation has at least as much taste for  $Q$ -informativeness induces the same order as  $\rho^{-1}$ .*

*Proof.* See appendix. □

*Remark.* Both axioms in Theorem 1.21 are essential: As the following examples demonstrate, omitting either admits indices to which the local taste is not ordinally equivalent to  $\rho^{-1}$ .

**Example 1.16.**  $Q \equiv 5$  satisfies local consistency, but it does not satisfy homogeneity of degree  $k < 0$ . The coefficient of local taste for this index induces the trivial order.

**Example 1.17.**  $Q := \frac{1}{\mu}$  satisfies homogeneity, but violates local consistency. The coefficient of local taste for this index induces the trivial order.

**Theorem 1.22.** (i) *Given  $k > 0$ , there exists a continuum of locally consistent homogeneous of degree  $-k$  indices of appeal for which the coefficient of local taste equals to the inverse of  $\rho$ .* (ii) *Moreover, some of these indices are not monotone with respect to Blackwell dominance.*<sup>54</sup>

---

<sup>54</sup> $Q$  is monotone with respect to Blackwell dominance if for any cost  $\mu > 0$  and all information structures  $\alpha, \beta$ , if  $\alpha$  Blackwell dominates  $\beta$  then  $Q(\mu, \alpha) > Q(\mu, \beta)$ .

*Proof.* See appendix. □

**Definition** ( $Q$ -informativeness globally more attractive). For an index  $Q$ , say that  $Q$ -informativeness *is globally at least as attractive* for  $u$  as it is for  $v$  (written  $v \lesssim_Q u$ ) if for all  $w, w'$ ,  $u$  at  $w$  has at least as much taste for  $Q$ -informativeness as  $v$  at  $w'$ .  $Q$ -informativeness *is globally more attractive* for  $u$  than to  $v$  (written  $v <_Q u$ ) if  $v \lesssim_Q u$  and not  $u \lesssim_Q v$ .

**Axiom** (Global consistency). For any  $w$ , any  $u, v$ , and any  $a, b \in \mathcal{A}$ , if  $v <_Q u$ ,  $A(a) < A(b)$  and  $v$  accepts  $a$  at  $w$ , then  $u$  accepts  $b$  at  $w$ .

**Theorem 1.23.** For a given  $k > 0$ ,  $A^k(\cdot)$  is the unique index that satisfies local consistency, global consistency and homogeneity of degree  $-k$  in prices, up to a multiplication by a positive number.

*Proof.* Let  $Q'$  satisfy the conditions and consider  $Q = (Q')^{1/k}$ . It is homogeneous of degree  $-1$  and still locally consistent, so by Theorem 1.21 the relation has at least as much taste for  $Q$ -informativeness induces the same order as  $\rho^{-1}$ . This, in turn, implies that if  $v \lesssim_Q u$  then  $v$  is uniformly more risk averse than  $u$ . Combined with this fact, global consistency and homogeneity of degree  $-1$  in prices imply the two axioms that are uniquely satisfied by positive multiples of  $A$ , according to Theorem 4 in Cabrales *et al.* (2012). That  $A$  satisfies local consistency follows from Theorem 1.20. This implies that  $A^k$  also satisfies local consistency. That other axioms are satisfied follows from Cabrales *et al.* (2012) using Theorem 1.20. The same holds for positive multiples of  $A^k$ . □

**Corollary 1.13.**  $J_e$ , the cost adjusted entropy reduction index, does not satisfy global consistency.

**Example 1.18.** (Based on Example 2 of Cabrales *et al.* (2012)). Let  $K = \{1, 2, 3\}$  and fix a uniform prior. Consider the information structures

$$\alpha_1 = \begin{bmatrix} 1 - \epsilon_1 & \epsilon_1 \\ 1 - \epsilon_1 & \epsilon_1 \\ \epsilon_1 & 1 - \epsilon_1 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 1 - \epsilon_2 & \epsilon_2 \\ 0.1 & 0.9 \\ \epsilon_2 & 1 - \epsilon_2 \end{bmatrix},$$

and the information transactions  $a_1 = (1, \alpha_1)$  and  $a_2 = (1, \alpha_2)$ . It can be shown that

$$A(a_1) \approx -\log\left(\frac{2}{3}\epsilon_1^{1/3} + \frac{1}{3}\epsilon_1^{2/3}\right),$$

and

$$A(a_2) \approx -\log\left(\epsilon_2^{1/3}\right).$$

This means that the ordering of the two transactions according to  $A$  depends on the choices of  $\epsilon_1, \epsilon_2 > 0$ . Even when they are both small, their relative magnitude matters.

In contrast, the cost adjusted entropy reduction index,  $J_e$ , ranks  $a_2$  higher than  $a_1$  for small  $\epsilon_1, \epsilon_2 > 0$ . To see this, note that

$$J_e(a_1) \approx \ln 3 - 0.462,$$

and

$$J_e(a_2) \approx \ln 3 - 0.550.$$

This means that there exists a choice of small enough  $\epsilon_1, \epsilon_2$  such that  $A(a_1) < A(a_2)$  and  $J_e(a_1) > J_e(a_2)$ . Hence, there exists two CARA functions with

different ARA coefficients (between  $A(a_1)$  and  $A(a_2)$ ), which both accept  $a_2$  but reject  $a_1$ , demonstrating that  $J_e$  violates global consistency.

As discussed previously, the homogeneity axiom has some cardinal content. In what follows, it will be removed and replaced with less demanding conditions: monotonicity in prices, and continuity with respect to prices. Example 1.19 will show that these conditions do not suffice to ensure that the local taste for  $Q$ -informativeness does not induce the trivial order or even that the index is monotonic with respect to Blackwell's order. As in previous sections, with a stronger version of global consistency, these conditions will suffice to pin down a unique index of informativeness (up to a monotonic transformation), and this index will have all of the desirable properties mentioned above.

**Definition** (Continuity). An index of informativeness is *continuous* (in price) if for every  $\alpha$ ,  $Q(\cdot, \alpha)$  is a continuous function from  $\mathbb{R}_+$  to  $\mathbb{R}_+$ .

**Example 1.19.**  $Q(\mu, \alpha) := 1 - \exp\left\{-\left(1 + \frac{1}{\mu}\right)\right\}$  is positive and continuous. It satisfies local consistency, but the relation has at least as much taste for  $Q$ -informativeness applies to any two utility-wealth pairs. Hence, for any  $u$  and  $v$ ,  $Q$ -informativeness is not more attractive for  $u$  than it is for  $v$ , and so global consistency is satisfied. The coefficient of local taste for  $Q$ -informativeness is equal to 1 for all agents at all wealth levels. Since  $Q$  is independent of the signal structure, it is clearly not monotonic with respect to Blackwell's order.

**Axiom** (Strong global consistency). *For any  $w$ , any  $u, v$ , and any  $a, b \in \mathcal{A}$ , if  $v \lesssim_Q u$ ,  $A(a) < A(b)$  and  $v$  accepts  $a$  at  $w$ , then  $u$  accepts  $b$  at  $w$ .*

Strong global consistency is clearly violated by the index from Example 1.19, as any two utilities  $u, v$  satisfy  $v \lesssim_Q u$ .

**Theorem 1.24.** *If  $Q$  is a continuous index of the appeal of information transactions that satisfies monotonicity in price and strong global consistency, and the relation has at least as much taste for  $Q$ -informativeness is reflexive, then  $Q$  is ordinally equivalent to  $A$ .*

*Proof.* See appendix. □

**Corollary 1.14.** *If  $Q$  is a continuous index of the appeal of information transactions that satisfies monotonicity in price and strong global consistency, and the relation has at least as much taste for  $Q$ -informativeness is reflexive, then  $Q$  satisfies monotonicity with respect to Blackwell dominance and local consistency, and the coefficient of local taste for  $Q$ -informativeness is ordinally equivalent to  $\rho^{-1}$ .*

### 1.9.3 The Demand for Information Transactions

**Proposition 1.13.** *An information transaction  $a$  with  $A(a) = b$  is rejected by  $u$  only if there exist some  $w$  such that local transactions with  $A$  of  $b$  are rejected. An information transaction  $a$  with  $A(a) = b$  is accepted by  $u$  only if there exist some  $w$  such that local transactions with  $A$  of  $b$  are accepted.*

*Proof.* Omitted. □

**Corollary.** *(Cabrales et al., 2014, Theorem 2) If  $A(a) > \sup_w \{\rho_u(w)\}$  then  $u$  rejects  $a$  at any wealth level. If  $A(a) < \inf_w \{\rho_u(w)\}$  then  $u$  accepts  $a$  at any wealth level.*



*Remark.* Cabrales *et al.* (2014) derive a result on *sequential transactions*,<sup>55</sup> which could be generalized to a result in the spirit of compound gamble property. Since this result requires some notation, I do not provide it here.

**Axiom** (Generalized Samuelson property).  $\forall u, w' \ S_Q^\infty(u, w') \leq \sup_w S_Q(u, w)$  and  $R_Q^\infty(u, w') \geq \inf_w R_Q(u, w)$ .

**Theorem 1.25.** *For a given  $k > 0$ ,  $A^k(\cdot)$  is the unique index that satisfies the generalized Samuelson property, local consistency and homogeneity of degree  $-k$  in prices, up to a multiplication by a positive number.*

*Proof.* Omitted. □

**Theorem 1.26.** *If  $Q$  satisfies the generalized Samuelson property, reflexivity, monotonicity with respect to first order time dominance and continuity then  $Q$  is ordinally equivalent to  $A$ .*

*Proof.* Omitted. □

#### 1.9.4 Properties of the Index $A$

The setting of information transactions is somewhat different than other settings that are discussed in this paper, in that the index depends on the prior, and is therefore not completely objective. Example 1.20 below shows that the order induced by  $A$  is different for different priors. Thus, the prior is a relevant part of the specification of the decision making problem that the index is derived from. The fact that in the setting presented here the prior

---

<sup>55</sup>Section 7.3 of Cabrales *et al.* (2014).

and the prices (which are more likely to be observable) coincide is comforting in this regard.<sup>56</sup>

An important property of the index  $A$  is that it is monotonic with respect to *Blackwell's (1953) partial ordering* of information structures (Cabrales *et al.*, 2014). According to Blackwell's order, one information structure is more informative than another if the latter is a garbling of the prior. Blackwell (1953) proved that one information structure is more informative than another according to this partial ordering if and only if every decision maker prefers it to the other. Cabrales *et al.* (2014) show that if  $\alpha$  is more informative than  $\beta$  in the sense of Blackwell, then  $A(\mu, \alpha) > A(\mu, \beta)$  for every  $\mu > 0$  and every prior.<sup>57</sup> As Blackwell's ordering is the parallel of stochastic dominance and time dominance, this property is analogous to the properties of the indices presented in previous sections. It is important to note that monotonicity with respect to Blackwell dominance was not one of the requirements in Theorem 1.24. Other desirable properties of the index include monotonicity in prices and being jointly continuous in  $p$ ,  $\mu$ , and  $q^s$ . For an extensive discussion of the properties of this index see Cabrales *et al.* (2014).

Finally, the cardinal interpretation of the index  $A$  is relatively more compelling, as the homogeneity (of degree -1) axiom may be interpreted as stating that the index measures information per dollar payed. If this interpretation is taken seriously, then the index may be used in practice for comparing different information providers, charging a fixed fee.

---

<sup>56</sup>See also the next subsection which discusses the prior-free implications of  $A$ .

<sup>57</sup>Recall that  $A$  depends on the prior  $p$ , even though this fact is not reflected in the notation I use.

### 1.9.5 Prior-Free Implications

In this section I make the dependence of  $A$  on the prior,  $p$ , explicit and write  $A(\cdot, p)$ . First, I note that the order induced by the index of the appeal of information transactions depends on the prior in the strict sense. This can be seen easily in the following example:

**Example 1.20.** Let  $K = \{1, 2, 3\}$  and let  $p_1 = (.5 - \epsilon, .5 - \epsilon, 2\epsilon)$  and  $p_2 = (2\epsilon, .5 - \epsilon, .5 - \epsilon)$ . Consider the information structures

$$\alpha_1 = \begin{bmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \\ .5 & .5 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} .5 & .5 \\ 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{bmatrix}$$

for some small  $\epsilon$ , and the information transactions  $a_1 = (1, \alpha_1)$  and  $a_2 = (1, \alpha_2)$ . It is easy to verify that  $A(a_1, p_1) > A(a_2, p_1)$ , but  $A(a_1, p_2) < A(a_2, p_2)$ . Informally, this is true since, given  $p_i$ ,  $\alpha_i$  reveals almost all of the information that an investor could hope for, but  $\alpha_{-i}$  could be improved upon significantly.

The upshot of the example is that without knowledge of the prior, an analyst cannot deduce the “correct” complete order which was derived previously. But some comparisons could still be made, even in the absence of knowledge about the prior. For example, since  $A$  is monotonic with respect to Blackwell dominance for all  $p$ , whenever one structure,  $\alpha$ , Blackwell dominated another,  $\beta$ , it is the case that  $A((\mu, \alpha), p) \geq A((\mu, \beta), p)$  for all prices,  $\mu$ , and all prior beliefs,  $p$ . The same holds for comparisons of structures that differ only in price.

**Definition.** An information transaction  $a$  is *at least as appealing as*  $b$  inde-

*p*endently of the prior if  $A(a, p) \geq A(b, p)$  for all prior beliefs,  $p$ .

As explained above, the order *prior-independent at least as appealing* is strictly partial, but it includes all the comparisons that could be made by Blackwell's partial order and monotonicity in prices. I now turn to show that it could compare strictly more pairs of information transactions. I base my proof on an example used in Shorrer (2015) to show that, even though the index  $I_e$  of Cabrales *et al.* (2013) depends on the prior, it can compare strictly more pairs of information structures than Blackwell's order.

**Theorem 1.27.** *There exists information transactions  $a = (1, \alpha)$  and  $b = (1, \beta)$ , such that  $\alpha$  does not dominate  $\beta$  in the Blackwell sense, yet  $a$  is at least as appealing as  $b$  independently of the prior, and  $b$  is not at least as appealing as  $a$  independently of the prior.*<sup>58</sup>

This result suggests that, even though the prior-independent order is partial, it still improves upon the more general Blackwell ordering. Thus, restricting attention to the particular decision making problem of investment, allows to derive a more complete order than Blackwell's, even without specifying a prior. This result, therefore, contributes to the literature which attempts to extend partial order of Blackwell by restricting the class of decision problems and agents under consideration (e.g. Persico, 2000; Athey and Levin, 2001; Jewitt, 2007).

*Proof.* Follows from the example. □

---

<sup>58</sup>Cabrales *et al.* (2014) disentangle the roles of  $p$ , and propose an index that depends on both security prices and the prior. The theorem will continue to hold in this setting, even if independence of both the prior and prices is required.

**Example 1.21.** Let  $K = \{1, 2\}$  and consider the information structures

$$\alpha_1 = \begin{bmatrix} .3 & .7 \\ .7 & .3 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} .3 & .7 \\ .1 & .9 \end{bmatrix},$$

and the transactions  $a = (1, \alpha_1)$ ,  $b = (1, \alpha_2)$ .

I claim that  $A(a, p) \geq A(b, p)$  for all  $p$ . Identify  $p$  with the probability of state 1, which lies in  $[0, 1]$ . Fixing the two information structures, define a function  $\phi_{a,b} : [0, 1] \rightarrow \mathbb{R}$  as follows:

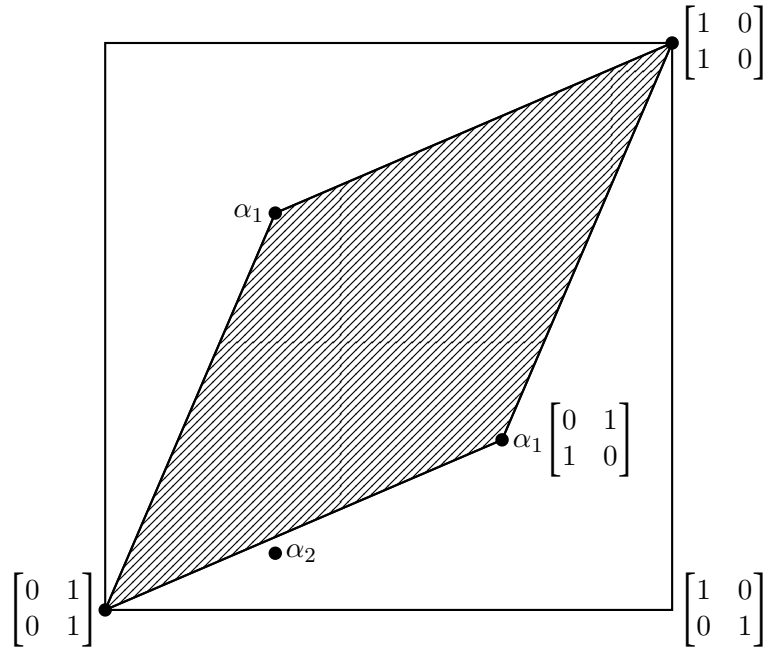
$$\phi_{a,b}(\cdot) := \exp\{-A(b, \cdot)\} - \exp\{-A(a, \cdot)\}.$$

For  $p \in \{0, 1\}$ ,  $A(\cdot, p) \equiv 0$ , hence  $\phi_{a,b}(p)$  also equals zero.  $\phi_{a,b}(\cdot)$  is also a continuous function (this follows from the properties of  $A$ ) and twice continuously differentiable in  $(0, 1)$  with a strictly positive second derivatives. This implies that  $\phi_{a,b}(\cdot)$  is a convex and continuous function with  $\phi_{a,b}(0) = \phi_{a,b}(1) = 0$ . But this means that  $\phi_{a,b}(p) \leq 0$  for all  $p \in [0, 1]$  which means that  $A(b, p) \leq A(a, p)$  for all  $p \in [0, 1]$ , hence  $a$  is at least as appealing as  $b$  independently of the prior. It is not hard to verify that  $b$  is not at least as appealing as  $a$  (by example, or using the strict convexity of  $\phi_{a,b}$ ).

Finally, it remains to check that the comparison is not due to monotonicity in price or Blackwell's order. The first is obvious, as  $a$  and  $b$  involve the same price. It is not very hard to verify that  $\alpha_1$  does not dominate  $\alpha_2$  in the Blackwell sense. To do this, note that the set of all  $2 \times 2$  information structures which are dominated by  $\alpha_1$  is

$$\text{Conv} \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} .3 & .7 \\ .7 & .3 \end{pmatrix}, \begin{pmatrix} .7 & .3 \\ .3 & .7 \end{pmatrix} \right\},$$

where Conv denotes the convex hull of the four matrices.  $\alpha_2$  is not included in this set, as Figure 1.1 illustrates.



**Figure 1.1:** The figure depicts the two dimensional space of  $2 \times 2$  information structures. These matrices could be written as  $\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$ , where both  $x$  and  $y$  are in  $[0,1]$ . In the figure,  $x$  is represented by the horizontal axis and  $y$  is represented by the vertical axis. The shaded area are the matrices which represent information structures which are dominated by  $\alpha_1$  in the Blackwell sense. The point  $\alpha_2$  is outside the shaded area.

## 1.10 Discussion

This paper presented an axiomatic approach for deriving an objective index which could serve as a guide for decision making for different decision makers. The approach was shown to pin down a unique index with desirable properties in five settings, demonstrating its generality and applicability for different decision making settings. This approach could potentially be used in other settings in which indices are needed. A particular setting which seems promising in this regard is the measurement of inequality, which has many similarities to the setting of risk (Atkinson, 1970). Future research should focus on characterizing the class of decision making problems to which the approach is applicable.

## Chapter 2

# More Complete Incomplete Orders of Informativeness and Segregation

### 2.1 Introduction

Understanding the demand for information is crucial for understanding many important economic environments. Yet, comparing the desirability of different information structures in a sensible way is an illusive undertaking. The reason is that in some settings certain pieces of information may be vital for some agents, while in other settings, or for different agents, other pieces of information will be more important.<sup>1</sup> The implication is that it is not

---

<sup>1</sup>The difficulty in generating a ranking which is independent of agents' preferences is discussed by Willinger (1989) in his paper which studies the relation between risk aversion and the value of information. Willinger (1989) discusses his choice of using the *expected value of information* (EVI) or "*asking price*" which was defined by LaValle (1968). The



possible to rank all information structures so that higher ranked structures are preferred to lower ranked ones by all agents at every decision making problem and for all prior beliefs. Some pairs of information structures may, however, be compared in this manner. In his seminal paper, Blackwell (1953) showed that one information structure is preferred to another by all agents in all settings if and only if the latter is a garbling of the former.<sup>2</sup> That is, if one is a noisy version of the other. But this order is partial and cannot be used to compare many pairs of information structures.<sup>3</sup>

Cabrales *et al.* (2013) use an approach of total rejections in the spirit of Hart (2011) to make such comparisons.<sup>4</sup> They restrict attention to investment decision problems, in a model a la Arrow (1972), and define a relation which they name *uniform investment dominance*. This turns out to be a complete order over all *information structures*, which refines the order suggested by Blackwell (1953). Their order, however, depends on the (unique, fixed, common) prior of the decision makers which are considered, and so they get, in fact, a continuum of orders, one for each prior. These orders are indeed different from each other; there exists pairs of information structures which are ranked differently depending on the prior selected. This means that *prior*

---

EVI measures a certain decision maker's willingness to pay for certain information, and so, "... the difficulty of defining a controversial continuous variable representing the 'amount of information' can be avoided."

<sup>2</sup>A simple proof is provided in Leshno and Spector (1992).

<sup>3</sup>Lehmann (1988); Persico (2000); Athey and Levin (2001); Jewitt (2007) and others, have extended this partial order by restricting the class of decision problems and agents under consideration.

<sup>4</sup>In fact, they follow Hart's *utility uniform rejections* which leads in his setting to the index suggested in Foster and Hart (2009). Hart (2011) Also suggested *wealth uniform rejections* which lead in his setting to the index suggested in Aumann and Serrano (2008). Cabrales *et al.* (2014) Later followed this second approach and suggested the index of appeal of information transactions.

*independent investment dominance* is a partial order.

The first part of this paper treats a question which was left unanswered in Cabrales *et al.* (2013): is prior independent investment dominance the same as Blackwell's partial order, or does it provide further insights for *prior-free* comparisons of information structures in *investment settings*. This question is important since an analyst cannot always observe the priors of agents in the market. My answer is that the latter is correct. I prove that there exist (many) pairs of information structures that *could* be compared by the partial order of prior independent investment dominance, and *cannot* be compared using Blackwell's order. I provide a complete characterization for these pairs of structures restricting attention to information structures with two states of the world and two signals.

In the second part of the paper, I turn to the measurement of *segregation*, another topic of interest for economists and other social scientists, which raises similar difficulties. Massey and Denton (1988) enumerate several dimensions of segregation. Much of the literature, including this paper, focuses on what they call *evenness*, the (dis)similarity of the distributions of different social groups among different locations. The standard model compares lists of locations and their group composition. Examples include measures of racial segregation where locations are physical locations and groups correspond to different ethnicities (Massey and Denton, 1988), and occupational gender segregation where locations correspond to different occupations and the groups correspond to genders (Flückiger and Silber, 1999).

One strand of the literature concentrates on partial orders. Duncan and Duncan (1955) treated the case of two groups (whites and non-whites), using

*segregation curves*. Segregation curves are analogous to Lorenz curves, where one race takes the role of the population, and the other takes the role of income. The members of the “population” who are located in locations with a lower proportion of the other race are treated as lower income individuals. Lasso de la Vega and Volij (2014) recently demonstrated the close connection of this literature to Blackwell’s order over information structures. They use an axiomatic approach to propose a (partial) order over cities according to the informativeness (in the sense of Blackwell) of neighborhoods about the ethnic groups of its residents. They show that for the case of two ethnic groups this order coincides with the one induced by segregation curves.<sup>5</sup>

Another strand of the literature is focused on complete orders and indices (see for example Massey and Denton, 1988; Flückiger and Silber, 1999; Rear-don and Firebaugh, 2002; Hutchens, 2004). Frankel and Volij (2011) present a sequence of ordinal axioms and show that different subsets of these axioms pin down different (classes of) indices, and that no complete order satisfies all of the axioms. One of the orderings they characterize is represented by the *mutual information* index, which is closely related to the index Cabrales *et al.* (2013) use to represent uniform investment dominance. I use this close relation and results from the first part of the paper to show that there exists a partial order that satisfies all of their axioms (except for completeness) and which strictly refines the partial orders of Lasso de la Vega and Volij (2014) and of Hutchens (2015).

---

<sup>5</sup>Hutchens (2015) later proposed a refinement of this order.

## 2.2 Preliminaries

I use the model and notation of Cabrales *et al.* (2013). I review it here briefly, as a complete discussion could be found in their paper.

I consider agents with a concave and twice continuously differentiable utility function for money, who have some initial wealth,  $w$ , and face uncertainty about the state of nature. There are  $K \in \mathbb{N}$  states of nature,<sup>6</sup>  $\{1, \dots, K\}$ , over which the agents have the prior  $p \in \Delta(K)$  which is assumed to have a full support.

I identify agents and utility functions, and denote the *Arrow-Pratt coefficient of relative risk aversion* of agent  $u$  at wealth  $w$  as (Pratt, 1964; Arrow, 1965, 1971):

$$\varrho_u(w) := -w \frac{u''(w)}{u'(w)}.$$

I restrict attention to agents with relative risk aversion that is increasing in their wealth (IRRA). This means that  $\varrho_u(\cdot)$  is non-decreasing for all agents considered. Justifications for this assumption include: theoretical considerations (Arrow, 1971), observed behavior in the field (Binswanger, 1981; Post *et al.*, 2008) and laboratory experiments (Holt and Laury, 2002). IRRA utility functions include *constant absolute risk aversion* (CARA) utilities, as well as *constant relative risk aversion* (CRRA) utilities (Hart, 2011). I further focus on agents that are *ruin averse*, namely, that satisfy  $\lim_{w \rightarrow 0^+} u(w) = -\infty$ . I denote by  $\mathcal{U}^*$  the class of these utility functions.

---

<sup>6</sup>With a slight abuse of notation, I also denote  $\{1, \dots, K\}$  by  $K$ . The meaning of  $K$  should be clear from the context.

The set of *investment opportunities*  $B^* = \left\{ b \in \mathbb{R}^K \mid \sum_{k \in K} p_k b_k \leq 0 \right\}$ , consists of all no arbitrage assets.<sup>7</sup> When an agent with initial wealth  $w$  chooses investment  $b \in B^*$  and state  $k$  is realized, his wealth becomes  $w + b_k$ . Hence,  $B^*$  includes in particular the option of inaction. The reference to the members of  $B^*$  as no arbitrage investment opportunities attributes to  $p_k$  an additional interpretation as the price of an Arrow-Debreu security that pays 1 if the state  $k$  is realized and nothing otherwise. Hence,  $p$  plays a dual role in this setting.

Agents may choose their investment, but I do not allow for bankruptcy (the possibility of negative wealth). I say that an investment  $b$  is *feasible* at wealth  $w$  when  $w + b_k \geq 0$  in every state  $k \in K$ . Before choosing a feasible investment, the agent has an opportunity to engage in an *information transaction*  $a = (\mu, \alpha)$ , where  $\mu > 0$  is the cost of the transactions, and  $\alpha$  is the information structure representing the information that  $a$  entails. To be more precise,  $\alpha$  is given by a finite set of signals  $S_\alpha$  and probability distributions  $\alpha_k \in \Delta(S_\alpha)$  for every  $k \in K$ . When the state of nature is  $k$ , the probability that the signal  $s$  is observed equals  $\alpha_k(s)$ . Thus, the information structure may be represented by a stochastic matrix  $M_\alpha$ , with  $K$  rows and  $|S_\alpha|$  columns, and the total probability of the signals is given by the vector  $p_\alpha := p \cdot M_\alpha$ . For simplicity, assume that  $p_\alpha(s) > 0$  for all  $s$ , so that each signal is observed with positive probability. Further, denote by  $q_k^s$  the probability that the agent assigns to state  $k$  conditional on observing the signal  $s$ , using Bayes' law. Note that although the notation does not indicate it,  $(q_k^s)_{k=1}^K = q^s \in \Delta(K)$

---

<sup>7</sup>I present a simplified version of Cabrales *et al.* (2013). Simplifications are for exposition purposes only, and have no effect on any of my results.

depends on  $\alpha$  and the prior  $p$ .

Agents are assumed to choose the optimal feasible investment opportunity in  $B^*$  given their belief,  $q$ . Therefore, the expected utility of an agent with utility  $u$ , initial wealth  $w$  and beliefs  $q$  is

$$V(u, w, q) = \sup_{b \in B^*, \text{feasible}} \sum_k q_k u(w + b_k).^8$$

In case that the agent acquires no information, his beliefs are given by the prior  $p$ . Since the agent is risk averse, in such case his optimal choice is inaction. So,

$$V(u, w, p) = u(w).$$

Accordingly, an agent *accepts* an information transaction if

$$\sum_s p_\alpha(s) V(u, w - \mu, q^s) > V(u, w, p) = u(w)$$

and *rejects* it otherwise.

The *entropy reduction* is defined by:

$$I(\alpha, p) = H(p) - \sum_s p_\alpha(s) \cdot H(q^s),$$

where,<sup>9</sup>

$$H(q) = - \sum_{k \in K} q_k \ln(q_k),$$

and  $x \ln x = 0$  by continuity.

---

<sup>8</sup>Throughout, I use the convention that  $(-\infty) \cdot 0 = 0$ .

<sup>9</sup>Note that  $q^s$  is not independent of the prior  $p$ , even though the dependence is not made explicit by the notation I use.

## 2.3 Results

**Definition.** For a fixed prior  $p$  information structure  $\alpha$  *uniformly investment-dominates* (or *investment dominates*, for short) information structure  $\beta$  whenever, for every wealth  $w$  and price  $\mu < w$  such that  $(\mu, \alpha)$  is rejected by all agents with utility  $u \in \mathcal{U}^*$  at wealth  $w$ ,  $\beta$  is also rejected by all those agents.

**Theorem.** [Cabrales, Gossner and Serrano] For a fixed prior  $p$ , information structure  $\alpha$  *investment-dominates* information structure  $\beta$  if and only if  $I(\alpha, p) \geq I(\beta, p)$ .

**Corollary.** If  $\alpha$  *Blackwell dominates*  $\beta$  then  $I(\alpha, p) \geq I(\beta, p)$  for all  $p$ .

**Definition.** An information structure  $\alpha$  *investment-dominates*  $\beta$  *independently of the prior* (or *prior-independently investment-dominates*), whenever  $\alpha$  *investment-dominates*  $\beta$  for any prior  $p$ .

**Theorem.** [Cabrales, Gossner and Serrano] There exists no linear ordering that orders information structures according to the ordering of investment dominance independently of the prior.

**Example 2.1.** Let  $K = \{1, 2, 3\}$  and let  $p_1 = (.5 - \epsilon, .5 - \epsilon, 2\epsilon)$  and  $p_2 = (2\epsilon, .5 - \epsilon, .5 - \epsilon)$ . Consider the information structures

$$\alpha_1 = \begin{bmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \\ .5 & .5 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} .5 & .5 \\ 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{bmatrix}.$$

It is easy to verify that  $I(\alpha_1, p_1) > I(\alpha_2, p_1)$ , but  $I(\alpha_1, p_2) < I(\alpha_2, p_2)$ . This is intuitive since, given  $p_i$ ,  $\alpha_i$  contains (almost) all the information that an investor could hope for, but  $\alpha_{-i}$  could be improved upon significantly.

While information structures cannot be linearly ordered according to investment dominance independently of the prior, the previous corollary suggests that this relation is not vacuous. There are some cases where one information structure investment-dominates another information for any prior, for example, when the former Blackwell-dominates the latter. A natural question that was left unanswered in Cabrales *et al.* (2013) is whether these are the only cases. Namely, are prior-independent investment-dominance and Blackwell-dominance the same? I answer this question in the negative.

**Theorem 2.1.** *There exists  $\alpha$  and  $\beta$  such that  $\alpha$  investment-dominates  $\beta$  independently of the prior, but  $\alpha$  does not dominate  $\beta$  according to Blackwell's order.*<sup>10</sup>

*Proof.* Follows from Example 2.2.<sup>11</sup> □

**Example 2.2.** Let  $K = \{1, 2\}$  and consider the information structures

$$\alpha_1 = \begin{bmatrix} .3 & .7 \\ .7 & .3 \end{bmatrix}, \alpha_2 = \begin{bmatrix} .3 & .7 \\ .1 & .9 \end{bmatrix}.$$

I claim that  $I(\alpha_1, p) \geq I(\alpha_2, p)$  for all  $p$ . I identify  $p$  with the probability of state 1, which lies in  $[0, 1]$ . Fixing the two information structures I define

---

<sup>10</sup>Using Example 2.2, Shorrer (2014) shows that the same applies to the index of appeal of information transactions (Cabrales *et al.*, 2014).

<sup>11</sup>I am grateful to Yufei Zhang for providing Example 2.2.



a function  $\phi_{\alpha_1, \alpha_2} : [0, 1] \rightarrow \mathbb{R}$  as follows:

$$\phi_{\alpha_1, \alpha_2}(\cdot) := I(\alpha_2, \cdot) - I(\alpha_1, \cdot).$$

For  $p \in \{0, 1\}$ ,  $I(\cdot, p) \equiv 0$ , hence  $\phi_{\alpha_1, \alpha_2}(0) = \phi_{\alpha_1, \alpha_2}(1) = 0$ . It is not hard to verify that  $\phi_{\alpha_1, \alpha_2}(\cdot)$  is a continuous function (this follows from the properties of  $I$ ).  $\phi_{\alpha_1, \alpha_2}(\cdot)$  is also twice continuously differentiable, and

$$\phi_{\alpha_1, \alpha_2}''(p) = \frac{0.0252 - 0.0192p}{(-0.3 - 0.4p)(0.7 - 0.4p)(0.3(-1 + p) - 0.1p)(1 + 0.3(-1 + p) - 0.1p)}. \quad (2.3.0.1)$$

This expression is always positive for  $p \in (0, 1)$ , which implies that  $\phi_{\alpha_1, \alpha_2}(\cdot)$  is a strictly convex and continuous function with  $\phi_{\alpha_1, \alpha_2}(0) = \phi_{\alpha_1, \alpha_2}(1) = 0$ . But this means that  $\phi_{\alpha_1, \alpha_2}(p) < 0$  for all  $p \in (0, 1)$  which means that  $I(\alpha_2, p) - I(\alpha_1, p) \leq 0$  for all  $p \in [0, 1]$ , hence  $\alpha_1$  *investment-dominates*  $\alpha_2$  *independently of the prior*.

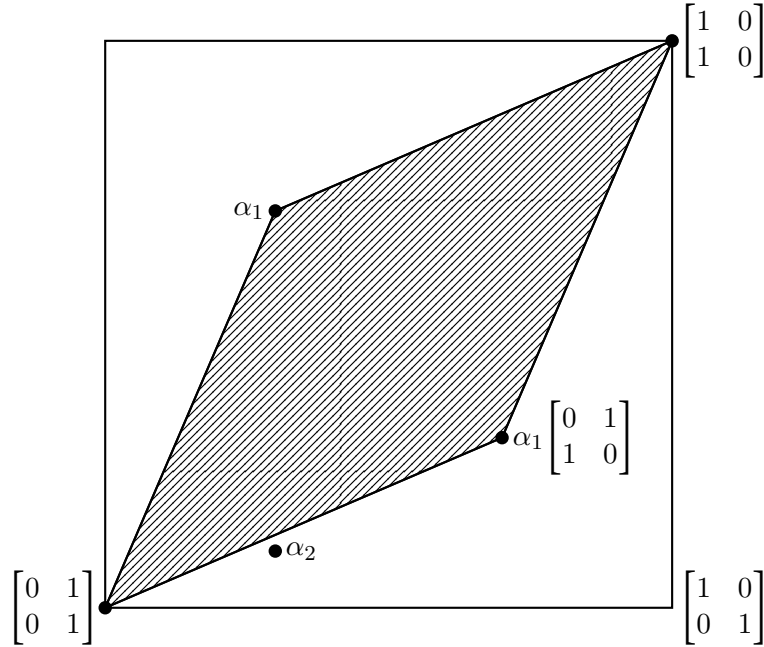
It remains to show that  $\alpha_1$  does not Blackwell-dominates  $\alpha_2$ . Let us look at the geometry of Blackwell dominance more generally. Given a  $K \times S$  information structure  $\alpha$ , the set of all  $K \times S$  information structures dominated by  $\alpha$  is defined as  $\text{Dom}(\alpha) = \{\alpha M : M \in (\Delta(S))^S\}$ . Namely,  $M$  ranges over all  $S \times S$  stochastic matrices. Being a linear image of the polytope  $(\Delta(S))^S$ ,  $\text{Dom}(\alpha)$  is a polytope whose vertices are images of the vertices of  $(\Delta(S))^S$ . Namely,

$$\text{Dom}(\alpha) = \text{conv} \left\{ \alpha \begin{bmatrix} e_{i_1} \\ \vdots \\ e_{i_S} \end{bmatrix} : e_{i_j} \text{ are vertices of } \Delta(S) \right\}.$$

The case of just two states ( $K = 2$ ) and two signals ( $S = 2$ ) is particularly simple. The set of all  $2 \times 2$  information structures dominated by a given  $2 \times 2$  information structure form a parallelogram,

$$\text{Dom}\left(\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}\right) = \text{conv}\left\{\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}, \begin{bmatrix} y & 1-y \\ x & 1-x \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right\}.$$

From Figure 2.1, it is easily seen that the information structure  $\alpha_2$  from the example is not Blackwell dominated by  $\alpha_1$ .



**Figure 2.1:** The figure depicts the two dimensional space of  $2 \times 2$  information structures. These matrices could be written as  $\begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$ , where both  $x$  and  $y$  are in  $[0,1]$ . In the figure,  $x$  is represented by the horizontal axis and  $y$  is represented by the vertical axis. The shaded area are the matrices which represent information structures which are dominated by  $\alpha_1$  in the Blackwell sense. The point  $\alpha_2$  is outside the shaded area.

The above counterexample extends to any number of states and signals. I now focus on the case of  $2 \times 2$  information structures (2 states of the world and 2 signals), and provide a complete characterization of comparable pairs.

**Definition 2.1.** Given two  $2 \times 2$  information structures  $\alpha$  and  $\beta$ , the function  $\phi_{\alpha,\beta} : [0, 1] \rightarrow \mathbb{R}$  is defined as follows:

$$\phi_{\alpha,\beta}(\cdot) := I(\beta, \cdot) - I(\alpha, \cdot).$$

**Theorem 2.2.** For  $2 \times 2$  information structures  $\alpha$  and  $\beta$ ,  $\alpha$  investment-dominates  $\beta$  independently of the prior if and only if

$$\phi'_{\alpha,\beta}(0^+) \leq 0, \text{ and}$$

$$\phi'_{\alpha,\beta}(1^-) \geq 0.$$

Furthermore,  $\alpha$  and  $\beta$  are comparable using this partial order if and only if

$$\phi'_{\alpha,\beta}(0^+) \phi'_{\alpha,\beta}(1^-) \leq 0.$$

*Proof.* By definition,  $\alpha$  investment-dominates  $\beta$  independently of the prior if and only if  $\phi_{\alpha,\beta}$  is non-positive on the interval  $[0, 1]$ . My proof is a generalization of the investigation of the function  $\phi_{\alpha_1,\alpha_2}$  in the proof of Theorem 2.1. The main step is to show that  $[0, 1]$  could be divided into two intervals,  $[0, t]$  and  $[t, 1]$  (with  $t$  possibly equal to 0 or 1), such that in one of these intervals  $\phi_{\alpha,\beta}$  is convex and the other is concave. In other words:  $\phi_{\alpha,\beta}$  is either convex, concave, or “*S-shaped*”: convex on one side of  $t$  and concave on the other side. From this analysis and the fact that  $\phi_{\alpha,\beta}(0) = \phi_{\alpha,\beta}(1) = 0$ , one readily concludes the first part of the theorem. The second part follows

from the first, since  $\phi_{\alpha,\beta} = -\phi_{\beta,\alpha}$ .

I turn now to proving that for every  $2 \times 2$  structures  $\alpha$  and  $\beta$ , there exist  $t \in [0, 1]$  such that  $\phi_{\alpha,\beta}$  is either convex or concave in each one of the intervals  $[0, t]$  and  $[t, 1]$  separately.

Denote

$$\alpha = \begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}, \quad \beta = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix},$$

for some  $x, y, a, b \in [0, 1]$ . The function  $\phi_{\alpha,\beta}$  is continuously twice differentiable on  $[0, 1]$ . Assume first that neither of the information structures is degenerate, that is, equals  $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$  or  $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ . Direct computation shows that

$$\phi''_{\alpha,\beta}(p) = \frac{(x-y)^2 p_\beta(1) p_\beta(2) - (a-b)^2 p_\alpha(1) p_\alpha(2)}{p_\alpha(1) p_\alpha(2) p_\beta(1) p_\beta(2)}. \quad (2.3.0.2)$$

The denominator is positive on  $(0, 1)$ , as a multiplication of four positive factors. The nominator is an affine function in  $p$ : it is the difference of two quadratic polynomials that have the same quadratic term. This implies that the entire expression could change signs at most once. This concludes the proof in the non-degenerate case.

If both  $\alpha$  and  $\beta$  are degenerate, then  $\phi_{\alpha,\beta} \equiv 0$ ; if only  $\alpha$  is degenerate, then

$$\phi''_{\alpha,\beta}(p) = \frac{-(a-b)^2}{p_\beta(1) p_\beta(2)};$$

and if only  $\beta$  is degenerate, then

$$\phi''_{\alpha,\beta}(p) = \frac{(x-y)^2}{p_\alpha(1) p_\alpha(2)}.$$

In all of the three degenerate cases  $\phi''_{\alpha,\beta}$  has the same sign throughout the interval  $(0,1)$ .  $\square$

The following theorem provides a sufficient condition which is much simpler to verify and illustrate than the condition of Theorem 2.2.

**Theorem 2.3.** *For a non-degenerate information structure  $\alpha = \begin{bmatrix} x & 1-x \\ y & 1-y \end{bmatrix}$*

*and an information structure  $\beta = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix}$ ,  $\alpha$  investment-dominates  $\beta$  independently of the prior if*

$$(x-y)^2 a(1-a) - (a-b)^2 x(1-x) \geq 0,$$

*and*

$$(x-y)^2 b(1-b) - (a-b)^2 y(1-y) \geq 0.$$

Note that Theorem 2.1 follows from Theorem 2.3. The condition in Theorem 2.3 specifies an intersection of two ellipses which is a strictly convex set; therefore any non-extreme point on the relative boundary of  $\text{dom}(\alpha)$  is an internal point of the set of information structures investment-dominated by  $\alpha$  independently of the prior.

*Proof.* (of Theorem 2.3) With the notation of the proof of Theorem 2.2 it is sufficient to show that  $\phi_{\alpha,\beta}$  is convex in  $[0,1]$ , or equivalently,  $\phi''_{\alpha,\beta}$  is non-negative in  $(0,1)$ . As seen through (2.3.0.2), since the denominator is always positive, it is sufficient to show that the nominator is non-negative.

Namely,

$$L(p) = (x - y)^2 p_\beta(1) p_\beta(2) - (a - b)^2 p_\alpha(1) p_\alpha(2) \geq 0 \quad \forall p \in (0, 1).$$

Since  $L(p)$  is a linear function of  $p$ , one needs only verify that the two end points, at  $p = 0, 1$ , are non-negative, which is exactly the condition of the theorem.  $\square$

*Remark 2.1.* Using this sufficient condition, one can show that two  $2 \times 2$  information structures drawn uniformly at random are comparable with probability greater than .84, compared with a  $2/3$  probability that they are comparable using Blackwell's criterion.

## 2.4 Segregation

In this section I use the previous results to show that all of the axioms suggested by Frankel and Volij (2011), with the exclusion of completeness, are satisfied simultaneously by a partial order which refines the ones of Lasso de la Vega and Volij (2014) and of Hutchens (2015). In particular, for the two ethnic groups case it refines the order induced by the segregation curves criterion.

### 2.4.1 Preliminaries

I use the model and notation of Frankel and Volij (2011). I review it here briefly, as a complete discussion could be found in their paper.

Population is assumed to be a continuum. I refer to groups as *ethnic groups* and to locations as *schools*. The objects to be ranked are (*school*)

*districts*, which are defined as follows:

**Definition.** A *district*  $X = \left\langle N, G, (T_g^n)_{g \in G, n \in N} \right\rangle \equiv \langle N(X), G(X), T(X) \rangle$  is a triplet where  $N$  is a finite non-empty set of *schools*,  $G$  is a finite non-empty set of *ethnic groups*, and for each  $g \in G$  and  $n \in N$ ,  $T_g^n$  is a non negative real number representing the number of members of ethnic group  $g$  attending school  $n$ .

When there is no risk of confusion, I will sometime just write the ethnic composition of the schools in the district. For example,  $\langle (1, 3), (2, 2) \rangle$  represents a district with two schools, and two ethnic groups (labeled black and white). The first has one black student and three white students, and the second has two students from each ethnic group.

For any scalar  $\alpha > 0$ ,  $\alpha X$  denotes a district in which for all  $n \in N$  and  $g \in G$  the number of students of ethnicity  $g$  in school  $n$  is multiplied by  $\alpha$ . For any two districts  $X, Y$  with the same ethnic groups  $X \uplus Y$  denotes their union into a single district. For example, if  $X = \langle (1, 2), (5, 5) \rangle$  and  $Y = \langle (3, 4) \rangle$  then  $2X = \langle (2, 4), (10, 10) \rangle$  and  $X \uplus Y = \langle (1, 2), (5, 5), (3, 4) \rangle$ .

Some more notation will prove useful.

$T_g = \sum_{n \in N} T_g^n$ : the number of students of ethnic group  $g$  in a district,

$T^n = \sum_{g \in G} T_g^n$ : the number of students attending school  $n$ ,

$T = \sum_{n \in N} \sum_{g \in G} T_g^n$ : the total number of students in the district,

$P_g = \frac{T_g}{T}$ : the proportion of group  $g$  students in the district,

$\pi^n = \frac{T^n}{T}$ : the fraction of the population attending school  $n$ ,

$p_g^n = \frac{T_g^n}{T^n}$  (when  $T^n > 0$ , and 0 otherwise): the proportion group  $g$  students

in school  $n$ ,

$t_g^n = \frac{T_g^n}{T_g}$  (when  $T_g > 0$ , and 0 otherwise): the fraction of group  $g$  students attending school  $n$ .

A general class of districts is denoted by  $\mathcal{C}$ . Sometimes, a class of districts with  $K$  ethnic groups will be considered. Such class will be denoted by  $\mathcal{C}_K$ . Additionally,  $\mathcal{C}^A = \bigcup_{K \geq 2} \mathcal{C}_K$ .

Given a class of districts  $\mathcal{C}$ , an *ordering of segregation*, or a *segregation ordering*, is a binary relation on that class which is transitive and reflexive. I denote such ordering by  $\succeq$  and interpret the statement  $X \succeq Y$  to mean “district  $X$  is at least as segregated as district  $Y$ .”  $>$  and  $\sim$  are derived in the usual way. Throughout attention will be restricted to orderings that treat schools symmetrically.<sup>12</sup>

The *Mutual Information index* is equal to the difference between the entropy of a district’s ethnic distribution and the weighted average entropy of the ethnic distributions of its schools:

$$M(X) = H(P) - \sum_{n \in N(X)} \pi^n H(p^n).$$

Restricting attention to districts with two ethnic groups  $\{W, B\}$ , it is possible to associate each district with a *Lorenz segregation curve*, the equivalent of a Lorenz curve where individuals from one race takes the role of the “population” and the role of “income” is taken by the proportion of members of the other race in their school. If the Lorenz segregation curve of  $X$  is always below that of  $Y$  we say that  $Y$  *at least as segregated as  $X$  according to the Lorenz criterion* (written  $Y \succeq_L X$ ). Lasso de la Vega and Volij (2014)

---

<sup>12</sup>That is, any permutation on the set of schools in  $X$  induces a district which is exactly as segregated.



show that  $Y \succeq_L X$  if and only if the information structure given by the matrix  $p(Y)$  (with  $p(Y)_{n,g} = p_g^n(Y)$ ) Blackwell dominates the one given by  $p(X)$ . Thus, the *Blackwell criterion* could be viewed as a generalization of the Lorenz criterion that is applicable even when more than two ethnic groups are present. Hutchens (2015) proposes a refinement of this orders which allows to permute the labels of the ethnic groups in one of the districts before applying the Lorenz criterion. I refer to his criteria as *Symmetric Lorenz* and *symmetric Blackwell* respectively.

### 2.4.2 Axioms

The following axioms are taken from Frankel and Volij (2011). As they discuss extensively, most of them also appear elsewhere in the segregation literature.

**Axiom** (Completeness (C)). *For all  $X, Y \in \mathcal{C}$ ,  $X \succeq Y$  or  $Y \succeq X$ .*<sup>13</sup>

**Axiom** (Nontriviality (N)). *There exists  $X, Y \in \mathcal{C}$ , such that  $X \succ Y$ .*

**Axiom** (Continuity (CONT)). *For any district  $Z \in \mathcal{C}$ , the set of districts that have the same groups and schools as  $Z$  and are at least as segregated as  $Z$  is closed, as is the set that have the same groups and schools as  $Z$  and are no more segregated than  $Z$ .*

**Axiom** (Scale Invariance (SI)). *For any district  $Z \in \mathcal{C}$  and any scalar  $\alpha > 1$ ,  $\alpha Z \sim Z$ .*

**Axiom** (Symmetry (SYM)). *The segregation ordering is invariant to permutations of the groups in the district. For any  $X = \langle N, G, T \rangle \in \mathcal{C}$  and any*

---

<sup>13</sup>The paper by Frankel and Volij (2011) only considers complete orders, and so this axiom is not stated explicitly.

permutation  $\sigma$  on  $G$ , the district  $X^\sigma = \langle N, G, \bar{T} \rangle$  such that  $\bar{T}_g^n = T_{\sigma(g)}^n$  satisfies  $X^\sigma \sim X$ .

**Axiom** (Independence (IND)). *Let  $X, Y \in \mathcal{C}$  have equal populations and equal group distributions. Then for any  $Z \in \mathcal{C}$ ,  $X \uplus Z \succeq Y \uplus Z$  if and only if  $X \succeq Y$ .*

**Axiom** (School Division Property (SDP)). *Let  $X \in \mathcal{C}$  be a district and  $n$  a school in  $X$ . Let  $X'$  be the district that results from splitting school  $n$  into two schools,  $n_1$  and  $n_2$ . Then  $X' \succeq X$ . Furthermore, if one of the new schools is empty, or the two schools have the same ethnic distribution ( $p^{n_1} = p^{n_2}$ ) then  $X' \sim X$ .*

**Axiom** (Composition Invariance (CI)). *For any district  $X \in \mathcal{C}$ , group  $g \in G(X)$ , and scalar  $\alpha > 0$ , let  $X'$  be the district resulting from multiplying the number of group  $g$  students in each school in  $X$  by  $\alpha$ . Then  $X' \sim X$ .*

**Axiom** (Group Division Property (GDP)). *Let  $X \in \mathcal{C}$  be a district in which the set of ethnic groups is  $G = G(X)$ . Let  $X'$  be the result of partitioning some group  $g \in G$  into two subgroups,  $g_1$  and  $g_2$ , such that either one subgroup is empty ( $T_{g_i} = 0$  for some  $i \in \{1, 2\}$ ) or the two subgroups have the same distribution across schools ( $t_{g_1}^n = t_{g_2}^n$  for all  $n \in N(X)$ ). Then  $X' \sim X$ .*

### 2.4.3 Results

**Theorem.** *[Frankel and Volij] An ordering on  $\mathcal{C}^A$  satisfies  $C$ ,  $SI$ ,  $IND$ ,  $SDP$ ,  $N$ ,  $GDP$ ,  $SYM$  and  $CONT$  if and only if it is represented by the Mutual Information index.*

**Corollary.** *[Frankel and Volij] No ordering on  $\mathcal{C}^A$  satisfies  $C$ ,  $SI$ ,  $IND$ ,  $SDP$ ,  $N$ ,  $GDP$ ,  $SYM$  and  $CONT$  and  $CI$ .*

*Proof.* The index  $M$  is not invariant to the transformation described in Axiom CI. □

**Example.** [Frankel and Volij] let  $X = \langle (10, 0), (0, 1000) \rangle$  and  $X' = \langle (1000, 0), (0, 1000) \rangle$ . CI requires that  $X \sim X'$ . but  $M(X) \neq M(X')$ , as the entropy of a district's ethnic distribution is greater under  $X'$  and the weighted average entropy of the ethnic distributions of schools is the same for both.

**Theorem.** *An order which satisfies SDP and SI is monotonic with respect to Blackwell's criterion.*

*Proof.* Frankel and Volij (2011) provide a proof assuming C, and their proof generalizes to this case as well. □

**Definition 2.2.**  $X \in \mathcal{C}^A$  is at least as segregated as  $Y \in \mathcal{C}^A$  according to the ordering  $\succeq^*$  if there exists a permutation matrix  $\Pi$  such that  $t(X)$  prior independently investment dominates  $\Pi t(Y)$ .

That is, I allow permuting the names of races in one of the districts before comparing the matrices  $t(X)$  and  $t(Y)$  using the prior independent investment dominance criterion.

**Theorem 2.4.** *The (partial) ordering on  $\mathcal{C}^A$   $\succeq^*$  strictly refines the symmetric Blackwell criterion and satisfies SI, IND, SDP, N, GDP, SYM and CONT and CI.*

The theorem says that  $\succeq^*$  satisfies all of the axioms stated above with the exclusion of completeness (C), and that it refines the orderings of Lasso de la Vega and Volij (2014) and Hutchens (2015).

*Proof.* Recall that  $M(X) = H(P) - \sum_{n \in N(X)} \pi^n H(p^n)$  satisfies all axioms but CI. Now, observe that  $I(t(X), P(X)) = H(P) - \sum_{n \in N} \pi^n \cdot H(p^n) = M(X)$ . It is not hard to verify that this implies that  $\succsim^*$  satisfies all the axioms but CI and C.

Furthermore, for any district  $X \in \mathcal{C}$ , group  $g \in G(X)$ , and scalar  $\alpha > 0$ , if  $X'$  is the district resulting from multiplying the number of group  $g$  students in each school in  $X$  by  $\alpha$ , then  $t(X) = t(X')$  (but  $P(X) \neq P(X')$ ). Thus,  $X \succsim^* Y$  if and only if  $X' \succsim^* Y$ . Hence,  $\succsim^*$  satisfies CI.

From Example 2.2 there exists  $\alpha$  and  $\beta$  such that  $I(\alpha, p) \geq I(\beta, p)$  for all  $p$ , and  $\alpha$  does not Blackwell dominates  $\beta$ , not even allowing for permutations as in the symmetric Blackwell criterion.<sup>14</sup> Thus, any two districts  $X$  and  $Y$  in  $\mathcal{C}^A$  with  $T(X) = T(Y) = 1$  and  $t(X) = \alpha$  and  $t(Y) = \beta$  have that  $X$  is more segregated than  $Y$  according to  $\succsim^*$  but not according to the symmetric Blackwell criterion.

Finally, the fact that the definition of  $\succsim^*$  allows for any permutation of  $t(Y)$ , combined with that fact that  $I$  is monotonic with respect to Blackwell dominance, implies that  $\succsim^*$  refines the symmetric Blackwell criterion.  $\square$

**Corollary 2.1.** *The restriction of  $\succsim^*$  to  $\mathcal{C}_2$  strictly refines the symmetric Lorenz ordering.*

---

<sup>14</sup>This part could be verified quickly using Theorem 2.3.

## Chapter 3

# Optimal Truncation in Matching Markets<sup>1</sup>

### 3.1 Introduction

One of the great success stories in economic theory is the application of matching theory to two-sided markets. A classic example is the National Resident Matching Program (NRMP), in which medical school students are matched to residency positions in hospitals. Rather than hospitals pursuing students via a decentralized series of offers, refusals and acceptances, matching occurs via a centralized mechanism. In this mechanism, each student ranks the hospital programs, and each hospital ranks the students. They submit these lists to an algorithm, which determines which students will be matched to which programs.

Such a centralized process has a number of advantages. First and foremost,

---

<sup>1</sup>Co-authored with Peter Coles

the algorithm on which this and many similar centralized processes are based produces an outcome that is *stable* with respect to reported preferences.<sup>2</sup> In a stable matching, no two agents mutually prefer each other to their assigned match, nor does any matched participant prefer to be unmatched. A second advantage is that eliminating a decentralized offer process may save time and other resources. Finally, as Roth and Xing (1994) have shown, a centralized mechanism can successfully halt the unraveling of a market.<sup>3</sup> Centralized matching mechanisms also power a variety of other markets, including the public school systems in New York, Boston, Singapore and other cities, as well as numerous specialized medical fellowships.

These centralized markets all employ versions of an algorithm proposed by Gale and Shapley (1962). The algorithm, which in one-to-one markets is often referred to as the *Men-Proposing Deferred Acceptance Algorithm* (*MP-DA*), takes as its inputs preference lists reported by agents, and outputs a stable matching. When agents are asked to report preference lists for submission to *MP-DA*, this begs the question: Do all agents have an incentive to report truthfully? Dubins and Freedman (1981) and Roth (1982) provide the answer: they do not. In fact, Roth showed that no mechanism that produces stable matchings will induce truth-telling as a dominant strategy for all agents. However, in the preference list submission game induced by

---

<sup>2</sup>In 1998, the algorithm used in the NRMP was altered to accommodate student couples and allow for specialized hospital positions, so that the outcome is “close to” a stable matching (see Roth and Peranson, 1999).

<sup>3</sup>Before the NRMP was introduced in the 1950s, offers and interviews were made as early as the fall of students’ third year in medical school, which was undesirable for a number of reasons. The willingness of both hospitals and students to participate argues strongly in favor of the program’s effectiveness. The NRMP enjoys participation rates of close to 100% of eligible students, with over 38,000 students participating in the March 2012 match.

*MP-DA*, for all participants on *one side* of the market, the “men,” truth-telling is a dominant strategy.<sup>4</sup> But this leaves open the question of how participants on the other side of the market, the “women,” might benefit by strategically misrepresenting their preferences.

Recent work has examined conditions under which gains to strategic manipulation are limited for all participants in the market, not just those on one side. One approach in the literature concerns *large markets*. Roth and Peranson (1999) observe that in the data from the NRMP, very few participants could have improved their outcomes by reporting different preferences. They show via simulations that when the length of preference lists is held fixed and the number of participants grows, the size of the set of stable matching shrinks (a property they term “core convergence”), so that opportunities for manipulation are reduced. Immorlica and Mahdian (2005) demonstrate this result theoretically, finding that in large marriage markets where preference list length is bounded, nearly all players have an incentive to truthfully report preferences. Kojima and Pathak (2009) generalize this result, showing that in many-to-one markets, preference list manipulation, as well as other modes of strategic manipulation such as non-truthful reporting of capacities (see also Sönmez, 1997), are again limited. Lee (2011) considers one-to-one matching markets where agent utilities are drawn from distributions with bounded support that have both a common and an independent component. He shows that when all agents report truthfully, the proportion of participants who can

---

<sup>4</sup>This is true in one-to-one, or “marriage” markets, where each agent has the capacity to match with at most one other agent. In many-to-one settings, e.g. students matching to hospitals, truth-telling is no longer a dominant strategy when, in the Deferred Acceptance Algorithm, the “hospitals” side makes the offers (see Roth, 1985).

achieve a significant utility gain from manipulation vanishes as the market grows large.<sup>5</sup>

Our approach takes a different tack. We do not require preference lists to be short, and ask: how should players optimally misrepresent preferences in markets that do not satisfy non-manipulability conditions? How “far” could optimal behavior be from truthfulness? We wish to study optimal manipulation, along with payoffs and market-wide welfare effects, and ask how strategic behavior and outcomes change as we vary market conditions.

The particular form of strategic misrepresentation we focus on is preference list *truncation*; that is, listing in order the first several partners from one’s true preference list, and identifying all other partners as unacceptable. Truncation has an intuitive logic: by listing less-preferred partners as unacceptable, the probability of being matched with these partners drops to zero. Agents using this strategy might hope that correspondingly, the likelihood of being matched to a partner who remains on the truncated list will go up. In the context of *MP-DA*, this intuition is confirmed: submitting a truncated preference list weakly increases the likelihood of being matched to some agent on the truncated list, regardless of beliefs about the lists other agents submit. But submitting a truncated preference list is a risky strategy. Limiting acceptable partners also increases the likelihood of ending up with no match. Analysis of this tradeoff is the crux of the results in this paper.

While always a method for weakly increasing the likelihood of matching

---

<sup>5</sup>In a very different approach, Featherstone and Mayefsky (2010) run lab experiments in  $5 \times 5$  marriage markets, and find that participants have trouble learning to find beneficial deviations under *MP-DA*, even if there are potential gains (though participants have more success in finding successful manipulation when facing “priority” mechanisms).



with better-ranked opponents, in some uncertain settings, truncation is optimal: Roth and Rothblum (1999) show that when agents' beliefs satisfy a form of symmetry termed " $\mathcal{M}$ -symmetry," they can do no better than to truncate. Ehlers (2004) demonstrated that this result holds under somewhat more general conditions.

Whether optimal or not, we analyze truncation, both in symmetric and general settings. We ask: to what degree should players truncate, if at all? (Note that submitting one's true preference list is also a form of truncation.) Can a participant realistically gain from truncation when she is extremely uncertain about what opponents might report? If players anticipate that others may be truncating, how does this affect their behavior? Do participants truncate in equilibrium? What are the welfare implications in a truncation equilibrium?

To evaluate the consequences of truncation, we first characterize the payoff from truncation for a woman with general beliefs over the preference lists other agents will submit in terms of the distributions of her most and least preferred achievable mates. In a market with  $N$  men and  $N$  women, when a woman believes submitted preferences of others are uniformly chosen from the set of all full-length preference lists, she may safely truncate a large fraction of her list with low risk of becoming unmatched. Further, as there is a large gap between the expected rank of the mate she receives from truthful revelation and her most preferred achievable mate (Pittel (1989) shows these asymptote to  $N/\log N$  and  $\log N$  respectively), truncation can lead to gains. The optimal degree of truncation can be significant. We demonstrate that in large, balanced, uniform markets the optimal level of

truncation approaches 100%.<sup>6</sup> That is, when there are many agents in the market, a woman optimally submits an extremely short list relative to her full preference list – in the limit, the fraction of men that she leaves on her list goes to zero.

The two sides of the market diverge in their tastes for truncation equilibria. Compared to the outcome from truthful reporting, women prefer any equilibrium in which they all use truncation strategies; for men the opposite is true. Furthermore, if there are two truncation equilibria that can be compared in the degree of truncation, women prefer the equilibrium in which they truncate more, while men prefer the equilibrium in which women truncate less. Under uniform preferences, we demonstrate the existence of a symmetric equilibrium where all women use the same truncation strategy. However, even under uniform preferences, asymmetric equilibria, where women vary in their degrees of truncation, may also exist. In such equilibria, and in contrast to the across-equilibria results, the women who truncate *least* are best off. Intuitively, while women benefit from truncation by other women, they would prefer not to bear the risk of truncation themselves.

Relaxing the uniform preferences assumption and returning to the environment where players have arbitrary beliefs, we examine comparative statics. We find that optimal truncation levels vary with risk preference: regardless of beliefs over reported opponent preferences, the less risk-averse a player, the more she should truncate.

---

<sup>6</sup>Recently, Ashlagi *et al.* (2013) have shown that the requirement that the number of men and women is balanced is crucial for the existence of the large gap in expected ranks. We discuss the robustness of our results to imbalances in the number of men and women in Chapter 4.

We then turn to correlation in players' preferences. The correlated preferences we consider are meant to capture the notion that in many settings, agents largely agree in their preferences over partners on the other side of the market, but that an individual's preferences may idiosyncratically depart from common opinion. We find that the higher the likelihood a participant places on opponents having preferences similar to her own, the less she should truncate. Our findings largely corroborate the simulation results of Roth and Peranson (1999), who find that when preferences are correlated, the set of stable matchings is small, and therefore the set of submitted preference lists that could lead to gains is minimal. An important difference between our correlation setting and Roth and Peranson's is that we consider incomplete information, where realized matchings may be unstable with respect to true preferences (even while stable with respect to reported preferences).

To place this analysis in context, several comments are in order. While, for the reasons stated earlier, we believe truncating the *bottom* of one's list is an intuitive manipulation in the preference list submission problem, in different environments eliminating *better-ranked* members from one's preference list might be a reasonable strategy. For example, in the job market for economists, departments may choose not to interview certain highly-accomplished candidates, reasoning that these candidates will receive offers they prefer more (see Coles *et al.*, 2010). In efforts to best use costly and scarce interview slots, departments may effectively "top-truncate" their preferences lists, focusing instead on candidates more likely to ultimately accept an offer. In general, it is when market frictions generate costs that this behavior arises. Lee and Schwarz (2012) consider a setting where information acquisition is costly, so

that firms prefer to interview workers who have a high likelihood of accepting (and likelihood is based on the number and identity of other firms interviewing a worker). Coles *et al.* (2013) consider a setting where workers can signal their preferences to firms, so that firms may choose not to make offers to better-ranked candidates, and instead make offers to candidates who have indicated likeliness to accept. In our paper, the analysis is performed after any costly information gathering has taken place, so these considerations do not arise.

The rest of the paper is organized as follows: Section 3.2 lays out the stable marriage setting and illustrates the fundamental tradeoff associated with truncation. In Section 3.3 we characterize the return to truncation, first for general beliefs, then in a uniform setting. In Section 3.4, we prove the existence of a truncation equilibrium in symmetric settings, and explore equilibrium welfare implications. Based on this existence result, we provide simulation evidence for a significant degree of truncation in equilibrium. Section 3.5 and Section 3.6 examine how truncation behavior relates to risk preferences and correlation of agent preferences, respectively. Section 3.7 concludes.

## 3.2 Matching Markets Background

We begin by setting out the basic model of matching. In contrast to some of the well-known matching papers, we approach the notion of preferences of participants from a cardinal rather than an ordinal perspective, which

allows us to discuss choice under uncertainty.<sup>7</sup> Note, however, that standard matching results involving ordinal preferences also apply as we may infer preference orderings from cardinal utilities.

### 3.2.1 Marriage Markets and Stability

A *marriage market* of size  $N$  consists of a triplet  $(\mathcal{M}, \mathcal{W}, u)$ , where  $\mathcal{M}$  is the set of men,  $\mathcal{W}$  is the set of women,  $|\mathcal{M}| = |\mathcal{W}| = N$ , and  $u = \prod_{i \in \mathcal{M} \cup \mathcal{W}} u_i$  is the profile of preferences for men and women.<sup>8</sup>

Preferences  $u_m : \mathcal{W} \cup \{m\} \rightarrow \mathbb{R}$  for man  $m \in \mathcal{M}$  are given by a von Neumann-Morgenstern utility function in which  $m$  derives utility  $u_m(w)$  from matching with woman  $w$  and  $u_m(m)$  from remaining single. For simplicity, we assume that  $u_m$  is one-to-one, so that  $m$  is never indifferent between any two certain options. Preferences  $u_w$  for woman  $w \in \mathcal{W}$  are defined similarly on  $\mathcal{M} \cup \{w\}$ .

As  $u_m$  is one-to-one,  $m$ 's preferences  $u_m$  induce a strict preference ordering  $P_m$  over  $\mathcal{W} \cup \{m\}$ . We refer to  $P_m$  as  $m$ 's *preference list*. For example, if  $N = 3$ ,  $u_m(w_1) > u_m(w_3) > u_m(w_2) > u_m(m)$  yields preference list  $(w_1, w_3, w_2, m)$ , meaning  $m$  prefers woman  $w_1$  to  $w_3$  to  $w_2$  to being single. Note that men may prefer bachelorhood over some of the women. For example,  $(w_1, w_3, m, w_2)$  indicates that  $m$  prefers  $w_1$  to  $w_3$  to remaining single to  $w_2$ . We say that man  $m$  finds  $w$  *acceptable* if  $m$  prefers  $w$  to remaining single. When convenient, we

---

<sup>7</sup>Some matching papers do manage to study choice under uncertainty even when agents have only ordinal preferences. For example, Ehlers and Massó (2007) and Roth and Rothblum (1999) use the related concepts of “Ordinal Bayesian Nash Equilibrium” and “ $P_w$ -stochastic dominance,” respectively.

<sup>8</sup>The assumption  $|\mathcal{M}| = |\mathcal{W}|$  was made for technical and notational convenience. However, it plays an important role in Theorem 4.1, as discussed in the Chapter 4.

list only a man's acceptable women. Preference lists for women are defined similarly, and we let  $P$  denote the profile of preference lists.

A *matching* is a pairing of men and women, so that each woman is assigned at most one man and each man at most one woman. Formally, a matching  $\mu$  is a mapping from  $\mathcal{M} \cup \mathcal{W}$  to  $\mathcal{M} \cup \mathcal{W}$  such that for every  $m \in \mathcal{M}$ ,  $\mu(m) \in \mathcal{W} \cup \{m\}$ , and for every  $w \in \mathcal{W}$ ,  $\mu(w) \in \mathcal{M} \cup \{w\}$ , and also for every  $m, w \in \mathcal{M} \cup \mathcal{W}$ ,  $\mu(m) = w$  if and only if  $\mu(w) = m$ . When  $\mu(x) = x$ , agent  $x$  is *single* or *unmatched* under matching  $\mu$ . For agents that are not single, we refer to  $\mu(m)$  as  $m$ 's wife and  $\mu(w)$  as  $w$ 's husband. The terms partner and mate are also used. In a matching, each agent cares only about his or her partner, and not about the partners of other agents, so that we may discuss agent preferences over matchings.

Given preferences, a matching is stable if no agent desires to leave his or her mate to remain single, and no pair of agents mutually desire to leave their mates and pair with each other. Formally, given a matching  $\mu$ , we say that it is *blocked* by  $(m, w)$  if  $m$  prefers  $w$  to  $\mu(m)$  and  $w$  prefers  $m$  to  $\mu(w)$ . A matching  $\mu$  is *individually rational* if for each  $x \in \mathcal{M} \cup \mathcal{W}$  with  $\mu(x) \neq x$ ,  $x$  finds  $\mu(x)$  acceptable. A matching  $\mu$  is *stable* if it is individually rational and is not blocked. In general, more than one stable matching may exist for given preferences.

Given preferences, a woman  $w$  is *achievable* for  $m$  if there is *some* stable matching  $\mu$  in which  $w = \mu(m)$ . Achievable mates of women are defined similarly.

### 3.2.2 The Men-Proposing Deferred Acceptance Algorithm

In their seminal 1962 paper, Gale and Shapley prove that in any marriage market there exists a stable matching. To demonstrate this result, they propose an algorithm – the Men-Proposing Deferred Acceptance Algorithm (*MP-DA*) – to generate a stable matching given any profile of preferences lists.

*MP-DA* takes as its input a preference list profile  $P$  for agents  $\mathcal{M} \cup \mathcal{W}$ , and the output is a matching  $\mu^M[P]$ . When  $P$  is clear from the context, we write  $\mu^M$  to denote  $\mu^M[P]$ .<sup>9</sup> The algorithm works iteratively as follows:

- **Step 1.** Each man proposes to the first woman on his preference list. Each woman then considers her offers, rejects all men deemed unacceptable, and if any others remain, rejects all but her most preferred mate.
- **Step  $k$ .** Each man who was rejected in step  $k - 1$  makes an offer to the next woman on his preference list. If his preference list is exhausted, or if he prefers bachelorhood to the next woman on his list, he makes no offer. Each woman behaves as in step 1, considering offers in hand (including any man she has retained from the previous step) and rejects all but her most preferred acceptable suitor.
- **Termination.** If in any step  $k$ , no man makes an offer, the algorithm terminates. Each woman is paired with her current mate and this matching is final.

---

<sup>9</sup>In one-to-one markets, the women-proposing version of the algorithm (*WP-DA*) has identical but reversed properties, with output denoted by  $\mu^W[P]$ .

Gale and Shapley show that this algorithm must terminate in finite time, and they provide a remarkable characteristic of the resulting outcome.

**Theorem.** (*Gale-Shapley*) *The matching  $\mu^M$  resulting from MP-DA is stable. Furthermore, for any other stable matching  $\mu$ , every man weakly prefers  $\mu^M$  to  $\mu$  and every woman weakly prefers  $\mu$  to  $\mu^M$ .*

The stability of the matching produced by *MP-DA* offers numerous advantages, as outlined in the introduction. But men are particularly satisfied with this outcome. For the men, the algorithm produces the *optimal* stable matching, based on reported preferences. For the women, however, this is not the case. As we will see, this feature may mean some women prefer to strategically misreport preferences, causing the algorithm to produce a different matching.

### 3.2.3 The Preference List Submission Problem for Men

We now turn to the incentive properties of *MP-DA*. That is, in a setting where agents are asked to submit preference lists to the algorithm, we ask if they have an incentive to report something other than the truth. We will see that women may, while men do not.

Consider a set of agents  $\mathcal{M} \cup \mathcal{W}$ . Agent  $i \in \mathcal{M} \cup \mathcal{W}$  with preferences  $u_i$  must submit a preference list  $\hat{P}_i$  to *MP-DA*, where  $\hat{P}_i$  is chosen from the set of  $i$ 's possible preference lists  $\mathcal{P}_i$ . The agent's beliefs about what preference lists others will report are represented by the random variable  $\tilde{P}_{-i}$ , which takes as its range  $\mathcal{P}_{-i}$ , the set of all possible preference list profiles for others. Note that since  $u_i$  is a von Neumann-Morgenstern utility function, agent  $i$



may compare outcomes in this incomplete information setting.

Agent  $i$  solves the *Preference List Submission Problem*:

$$\max_{\hat{P}_i \in \mathcal{P}_i} \mathbb{E}[u_i(\mu^M[\hat{P}_i, \tilde{P}_{-i}](i))].$$

Dubins and Freedman (1981) and Roth (1982) have shown that for any man  $m$  with preferences  $u_m$  and beliefs  $\tilde{P}_{-m}$ , it is optimal for  $m$  to submit his true preference list  $P_m$  (which corresponds to  $u_m$ ).

**Theorem.** (*Dubins and Freedman; Roth*) *In the Preference List Submission Problem,*

$$P_m \in \arg \max_{\hat{P}_m \in \mathcal{P}_m} \mathbb{E}[u_m(\mu^M[\hat{P}_m, \tilde{P}_{-m}](m))].$$

### 3.2.4 The Preference List Submission Problem for Women

For women submitting preference lists to *MP-DA*, truth-telling may not be optimal. One way a woman  $w$  might misrepresent preferences is by submitting a *truncation* of her true preference list; that is, listing in order the first several men from her true preference list and declaring all other men unacceptable. Truncation generates a tradeoff: it may cause a woman to match with the better-ranked men she leaves on her list, but may also cause her to be left unmatched. In this section we demonstrate this tradeoff, pose the problem of optimal truncation, and describe conditions so that in the Preference List Submission Problem, among all possible preference list submission strategies, truncation is optimal.

The following example demonstrates the tradeoff at hand.

**Example 3.1. Strategic Truncation.** Suppose men and women have the following preference lists:

$m_1$	$m_2$	$m_3$	$w_1$	$w_2$	$w_3$
<b>w<sub>3</sub></b>	<b>w<sub>2</sub></b>	<b>w<sub>1</sub></b>	$m_1$	$m_1$	<b>m<sub>1</sub></b>
$w_1$	$w_1$	$w_2$	$m_2$	$m_3$	$m_2$
$w_2$	$w_3$	$w_3$	<b>m<sub>3</sub></b>	<b>m<sub>2</sub></b>	$m_3$

We consider the strategic incentives of woman 1, assuming all other agents report truthfully. First, suppose  $w_1$  submits her true preference list. In this case, *MP-DA* stops after one step and  $w_1$  is matched to  $m_3$ , her least preferred mate. The stable matching is indicated above in bold.

Now suppose that  $w_1$  misrepresents her preferences and submits the truncated list  $(m_1, m_2)$ . In this case, she will reject man  $m_3$ 's first round offer in the *MP-DA*. Man  $m_3$  must then make an offer to  $w_2$  in the next round. Woman  $w_2$  will accept  $m_3$  over  $m_2$ , who made her an offer in the previous round. Man  $m_2$  then finds himself single, and must make an offer to  $w_1$ . Woman  $w_1$  accepts  $m_2$ 's offer and *MP-DA* terminates, yielding the matching in bold below:

$m_1$	$m_2$	$m_3$	$w_1$	$w_2$	$w_3$
<b>w<sub>3</sub></b>	$w_2$	$w_1$	$m_1$	$m_1$	<b>m<sub>1</sub></b>
$w_1$	<b>w<sub>1</sub></b>	<b>w<sub>2</sub></b>	<b>m<sub>2</sub></b>	<b>m<sub>3</sub></b>	$m_2$
$w_2$	$w_3$	$w_3$		$m_2$	$m_3$

Therefore, by truncating her list,  $w_1$  improves her outcome.

To see how truncation can be dangerous, suppose  $w_1$  truncates her list

even more and submits  $(m_1)$  only. In this case, the algorithm will leave her unmatched, as shown in bold below:

$m_1$	$m_2$	$m_3$	$w_1$	$w_2$	$w_3$
<b><math>w_3</math></b>	$w_2$	$w_1$	$m_1$	$m_1$	<b><math>m_1</math></b>
$w_1$	$w_1$	<b><math>w_2</math></b>		<b><math>m_3</math></b>	$m_2$
$w_2$	$w_3$	$w_3$		$m_2$	$m_3$

Remark 3.1 characterizes woman  $w$ 's match when she submits a truncated version of her preference list, demonstrating generally how truncation can lead to the three outcomes in our example. For  $k \in \{0, \dots, N\}$ , we denote by  $P_w^k$  the preference list which includes in order only  $w$ 's  $k$  most preferred men, and call this the  $k$ -truncation of her true preference list  $P_w$ . If fewer than  $k$  men are acceptable to  $w$ , then  $P_w^k \equiv P_w$ .

*Remark 3.1.* Let  $P$  be the preference list profile of all agents in  $\mathcal{M} \cup \mathcal{W}$ . Then  $\mu^M[P_w^k, P_{-w}](w)$  is  $w$ 's least preferred achievable mate under  $P$  with rank  $\leq k$ . Should no such mate exist,  $\mu^M[P_w^k, P_{-w}](w) = w$ .

The example illustrates a general principle; given the preference lists submitted by others, truncation by woman  $w$  can have one of three consequences:

1. No effect. Woman  $w$  has truncated below her least preferred achievable mate.
2. Improvement. Woman  $w$  truncates above her least preferred mate, and is matched with her least preferred achievable mate above the point of truncation.
3. Unmatched. Woman  $w$  has over-truncated, truncating above her most

preferred achievable mate.

If woman  $w$  is certain of the preference lists  $P_{-w}$  others are submitting, her truncation decision is simple: she calculates her most preferred achievable mate under  $P = (P_w, P_{-w})$  and truncates her list to just include him. If instead  $w$  believes her opponents will submit preference lists according to some probability distribution, then truncating her list at  $k$  generates a lottery over outcomes in which either her partner will be among her  $k$  most preferred men, or else she will be unmatched. This tradeoff, between improvement and becoming unmatched, will guide our analysis in this paper.

### Optimality of Truncation

Truncation is not the only possible misrepresentation of preferences. A woman could reverse two men in her preference list, list men as acceptable who are in fact unacceptable, drop men from the middle of her list, or use some combination of these. However, under some conditions, truncation is optimal.

The next proposition states that under certainty, women can do no better than to truncate (Roth and Vande Vate, 1991).

**Proposition.** *(Roth and Vande Vate) Suppose woman  $w$  has preferences  $u_w$  and knows others will report preference lists  $P_{-w}$  to MP-DA. Then truncating such that  $\mu^W(w)$  is the last acceptable partner on her list is an optimal strategy for  $w$ .*

Perhaps surprisingly, when a woman has very little information about the preference lists others might report, she again can do no better than to truncate. In order to gain from non-truncation misrepresentations, such as

swapping the positions of two men in her reported preference list, a woman must have very specific information about the preference lists others report. Without such information, it is best to leave the men in their correct order. Roth and Rothblum (1999) demonstrate this principle using the following framework.

Let woman  $w$ 's beliefs about reported preference lists of others be represented by  $\tilde{P}_{-w}$ , a random variable taking on values in  $\mathcal{P}_{-w}$ . If  $P_{-w}$  is a preference list profile for agents  $-w$ , define  $P_{-w}^{m \leftrightarrow m'}$  to be the preference list profile in which  $m$  and  $m'$  swap preference lists, and all women swap the positions of  $m$  and  $m'$  in their lists. We say that woman  $w$ 's beliefs are  $(m, m')$ -symmetric if  $\Pr(\tilde{P}_{-w} = P_{-w}) = \Pr(\tilde{P}_{-w} = P_{-w}^{m \leftrightarrow m'})$  for all  $P_{-w} \in \mathcal{P}_{-w}$ . For a subset  $\mathcal{M}' \subseteq \mathcal{M}$ , beliefs  $\tilde{P}_{-w}$  are  $\mathcal{M}'$ -symmetric if they are  $(m, m')$ -symmetric for all  $m, m' \in \mathcal{M}'$ .

**Theorem.** (*Roth and Rothblum*) Suppose  $w$ 's beliefs about reported preference lists of others are  $\mathcal{M}$ -symmetric. Then any preference list  $\hat{P}_w$  she might submit to MP-DA is weakly  $P_w$ -stochastically dominated by some truncation of her true preference list.<sup>10</sup>

Hence, when  $w$  is certain about reported preference lists of her opponents, or when she has extreme, symmetric uncertainty, truncation is optimal.

---

<sup>10</sup>  $\hat{P}_w$  is  $P_w$ -stochastically dominated by  $\hat{P}'_w$  iff for any vNM utility function that corresponds to  $P_w$ , the expected utility from submitting  $\hat{P}'_w$  is at least as great as the expected utility from submitting  $\hat{P}_w$ .

## The Truncation Problem

Even when truncation is not optimal, we may sometimes wish to restrict the choice set for women to truncations of her true preference list. We define the *Truncation Problem* for woman  $w$  with preferences  $u_w$  and beliefs  $\tilde{P}_{-w}$  on others' submitted preference lists as

$$\max_{k \in \{0, \dots, N\}} \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))].$$

For convenience, whenever we consider the Truncation Problem for a woman  $w$ , we will relabel men so that  $w$  has  $u_w(m_1) > u_w(m_2) > \dots > u_w(m_N)$ .

## 3.3 Characterizing Truncation Payoffs

In this section we explore a woman's payoff from submitting a truncation of her true preference list. We first derive a formula for her payoff from truncating at any point in her list in terms of the distribution of her most and least preferred achievable mates. When a woman believes that reported preference lists of her opponents are distributed symmetrically over the set of all preference lists, we can say more about her gains from truncation: conditional on truncation yielding an improvement,  $w$ 's expected partner rank will be exactly half of  $(1 + \text{her point of truncation})$ . Further, in uniform markets, it is highly likely that for an individual woman, some degree of truncation will yield an improvement, and that, in fact, she may safely and beneficially truncate a large fraction of her list. We demonstrate that as the market grows large, the length of a woman's optimal reported list, as a fraction of her full preference list length, goes to zero. Hence, even in a setting

where agents possess very little information about opponent preferences, there is room for significant strategic misrepresentation.

### 3.3.1 Truncation in Two Stages: Match, then Divorce

To aid us in our analysis, we show that when woman  $w$  submits a  $k$ -truncation of her preference list to *MP-DA*, the outcome is identical to that from a two stage *Divorcing Algorithm*. In the first stage of the algorithm, agents submit preference lists to *MP-DA*. In the second stage,  $w$  ‘divorces’ her mate and declares all men with rank  $\geq k$  unacceptable. This sets off a chain of new offers and proposals, ending in a new match.<sup>11</sup>

The Divorcing Algorithm takes as its input a set  $P$  of preference lists, a woman  $w$ , and a truncation point  $k \in \{0, \dots, N\}$ , and generates a matching  $\mu^{DIV}[P, k, w]$ .

#### The Divorcing Algorithm

- **Step 0.** Initialization. Run *MP-DA* to find the men-optimal matching  $\mu^M[P]$ . If  $w$  is single or if  $w$ ’s mate has rank  $\leq k$  in  $P_w$ , terminate. Otherwise, divorce  $w$  from her mate. Declare candidates with rank  $\geq k$  unacceptable for  $w$ .

Iteration over steps 1 and 2:

- **Step 1.** Pick an arbitrary single man who has not exhausted his preference list. If no such man exists, terminate.

---

<sup>11</sup>Our Divorcing Algorithm is closely related to the techniques used in McVitie and Wilson (1971), where a “breakmarriage” operation is used to generate all the stable matchings for a given marriage market.

- **Step 2.** The man chosen in the previous step makes an offer to the most preferred woman on his preference list who has not already rejected him. If this woman finds the man acceptable and she prefers him to her current mate (or is single), she divorces (if necessary) and holds the new man's offer. Return to step 1.

The Divorcing Algorithm yields a matching identical to the output from  $w$ 's submission of a  $k$ -truncated list to  $MP\text{-}DA$ :

**Proposition 3.1.** *For all  $k \in \{0, \dots, N\}$ ,  $P \in \mathcal{P}$  and  $w \in \mathcal{W}$ , we have  $\mu^{DIV}[P, k, w] = \mu^M[P_w^k, P_{-w}]$ .*

With this equivalence in hand, when we consider the submission of a truncated preference list, we can think of it as a two stage process, focusing on the chain of offers (if there is one) in the Divorcing Algorithm. We will be interested in whether a chain will end with *i*) a new acceptable man proposing to woman  $w$ , or with *ii*) a single man making an acceptable offer to a single woman in  $\mathcal{W} \setminus w$ , or else exhausting his list. These outcomes correspond to truncation yielding an improvement over truthful reporting, and truncation leaving  $w$  unmatched, respectively. Knowing that following a “divorce,”  $w$  will receive at most one more offer will enable us to calculate the returns to truncation, conditional on truncation yielding an improvement.

### 3.3.2 Truncation under General Beliefs

In this section, we characterize woman  $w$ 's payoff from submitting a truncated version of her true preference list in terms of the distributions of her most and least preferred achievable mates. The results build on Remark 3.1, which



illustrates how in settings of certainty, a woman may gain, lose or see no change from truncation.

We consider the Truncation Problem for woman  $w$  with preferences  $u_w$  and beliefs  $\tilde{P}_{-w}$  about reports of other agents. Throughout the section,  $u_w$  (and hence,  $P_w$ ) is fixed, so we can denote  $w$ 's payoff from  $k$ -truncation when others submit preference lists  $P_{-w}$  as

$$v(k, P_{-w}) \equiv u_w(\mu^M[P_w^k, P_{-w}](w)).$$

Note that  $v(N, P_{-w})$  gives  $w$ 's payoff if she reports truthfully, and  $v(k, P_{-w}) = u_w(w)$  if  $k$ -truncation leaves  $w$  unmatched. The Truncation Problem then becomes

$$\max_{k \in \{0, \dots, N\}} \mathbb{E}[v(k, \tilde{P}_{-w})].$$

To evaluate  $\mathbb{E}[v(k, \tilde{P}_{-w})]$ , we condition on the three possible effects of truncation: no effect, improvement, and causing  $w$  to become unmatched.

Define  $k_l(P_{-w})$  and  $k_h(P_{-w})$  to be  $w$ 's rank of her mate under  $\mu^M[P]$  and  $\mu^W[P]$ , respectively. That is,  $k_l(P_{-w})$  ( $k_h(P_{-w})$ ) gives the rank of  $w$ 's least (most) preferred achievable mate when  $-w$  report preference lists  $P_{-w}$ . Set  $k_l(P_{-w}) = k_h(P_{-w}) = N + 1$  when  $w$  has no achievable mate. Let  $f(\cdot)$  be the probability mass function of the random variable  $k_l(\tilde{P}_{-w})$  so that

$$f(x) = \Pr(k_l(\tilde{P}_{-w}) = x)$$

for  $x \in \{1, \dots, N + 1\}$ . Let  $F(\cdot)$  be the associated cumulative distribution function. Similarly, let  $g(\cdot)$  be the probability mass function and  $G(\cdot)$  be the cumulative distribution function of the random variable  $k_h(\tilde{P}_{-w})$ .

Using  $F(\cdot)$ ,  $G(\cdot)$ , and Remark 3.1, we can express  $w$ 's expected payoff from  $k$ -truncating her list by using the law of conditional expectations:

$$\begin{aligned}\mathbb{E}[v(k, \tilde{P}_{-w})] &= F(k) \cdot \sum_{i=1}^k \frac{f(i)}{F(k)} u_w(m_i) \\ &+ [G(k) - F(k)] \cdot \mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)] \\ &+ [1 - G(k)] \cdot u_w(w),\end{aligned}\tag{3.3.2.1}$$

where the set  $\mathcal{P}_2(k) \equiv \{P_{-w} \mid v(k, P_{-w}) > v(N, P_{-w})\}$  gives the cases when truncation yields an improvement, compared to truthful reporting. When truncation causes  $w$  to be unmatched, her payoff is clearly  $u_w(w)$ , and when truncation has no effect, the likelihood of being matched with  $x$  is  $f(x)/F(x)$ .<sup>12</sup>

In the next two sections, we will focus on the middle term of the sum in 3.3.2.1; that is, the cases where truncation yields improvement. We will first see that when there are gains, the improvement can be significant. We will see later that in large markets, these gains may outweigh the risk of being left unmatched.

### 3.3.3 Truncation under $\mathcal{M}$ -Symmetric Beliefs

In this section, we examine the Truncation Problem when woman  $w$  has  $\mathcal{M}$ -symmetric beliefs. We show that conditional on  $w$ 's truncation yielding an improvement compared to truthful reporting, her mate is equally likely to be any man she lists as acceptable. This is somewhat surprising, because when  $w$  has unconditional  $\mathcal{M}$ -symmetric beliefs and submits preferences in the *MP-DA*, we would certainly not expect  $w$ 's mate to be uniformly distributed;

---

<sup>12</sup>If  $F(k) = 0$ , the first term in the sum is zero.

because of the stability requirement, she is far more likely to be matched with her more preferred mates.

**Lemma.** (*Truncation under  $\mathcal{M}$ -Symmetric Beliefs*) Suppose woman  $w$ 's beliefs  $\tilde{P}_{-w}$  about the reported preference lists of her opponents are  $\mathcal{M}$ -symmetric. Then according to her beliefs,

$$Pr\{\mu^M[P_w^k, \tilde{P}_{-w}](w) = m_i \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)\} = Pr\{\mu^M[P_w^k, \tilde{P}_{-w}](w) = m_j \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)\}$$

for all  $k \in \{1, \dots, N\}$ ,  $i, j \in \{1, \dots, k\}$ . Hence,

$$\mathbb{E}[v(k, \tilde{P}_{-w}) \mid \tilde{P}_{-w} \in \mathcal{P}_2(k)] = \frac{\sum_{i=1}^k u_w(m_i)}{k}.$$

The intuition in this result comes from the Divorcing Algorithm. Consider the settings where  $k$ -truncation will yield an improvement for  $w$  ( $P_{-w} \in \mathcal{P}_2(k)$ ). By first reporting her true preferences and then divorcing her mate, we know that there will ensue a chain of offers. This chain ends when exactly one man – ranked better than her former mate – will make an offer to  $w$ . By the symmetry of  $w$ 's beliefs, this is equally likely to be any of these men.

Crucial to the reasoning is that since we know truncation will yield an improvement, this corresponds to an algorithmic outcome where exactly one new superior offer is made to  $w$ . In *MP-DA* generally, multiple offers may be made to  $w$ , making it difficult to pinpoint the distribution of her mate's rank.<sup>13</sup>

---

<sup>13</sup>This result is related to the Principle of Deferred Decisions (“don’t do today what you can put off until tomorrow”), which was applied to the stable marriage problem in Knuth (1976). We may think of woman  $w$  as deferring her views on her preferences over men  $\{1, \dots, k\}$  until she actually receives an offer from one of them. When the first one arrives, only then does she assign the man a rank, which in expectation will be  $\frac{1+k}{2}$ .

With our Lemma in hand, we can now express  $w$ 's expected payoff from truncation at  $k$  as

$$\begin{aligned}
\mathbb{E}[v(k, \tilde{P}_{-w})] &= F(k) && \cdot \sum_{i=1}^k \frac{f(i)}{F(k)} u_w(m_i) \\
&+ [G(k) - F(k)] && \cdot \frac{\sum_{i=1}^k u_w(m_i)}{k} \\
&+ [1 - G(k)] && \cdot u_w(w).
\end{aligned} \tag{3.3.3.1}$$

### 3.3.4 Optimal Truncation in Large Markets

We now investigate optimal truncation for women when the market size grows large. We will focus on the special case of *uniform beliefs*. That is, when facing the Truncation Problem,  $w$  believes reported preference lists  $P_{-w}$  of her opponents to be chosen uniformly and randomly from the set of all possible full preference list profiles  $\mathcal{P}_{-w}$  (where a full preference list profile is a profile in which each agent prefers any possible mate to being unmatched). Uniform beliefs are a special case of  $\mathcal{M}$ -symmetric beliefs.<sup>14</sup> Hence, under uniform beliefs, we can be sure that truncation is optimal.

The stable marriage problem under uniform beliefs has received attention, especially in the mathematics and computer science literature, in large part because this setting facilitates average and worst case analyses (see Dzierzawa and Oméro, 2000; Knuth, 1976; Pittel, 1989). But for our purposes, uniform beliefs offer a tractable incomplete information setting where agents know little about the preferences of others.

---

<sup>14</sup>But uniformity is not equivalent to  $\mathcal{M}$ -symmetry. Under  $\mathcal{M}$ -symmetric beliefs, a woman may have specific knowledge about how the men rank *her*. For example, she may know that all the men prefer her to  $w_2$ . With uniform beliefs, such knowledge is ruled out.

Suppose that woman  $w$  has preferences  $u_w(\cdot)$  linear in the rank of her match (where being unmatched is treated as rank  $N + 1$ ), or else has any strictly increasing, convex transformation of such preferences. Suppose further that  $w$  has uniform beliefs. Define

$$k^*(N) \equiv \max \left( \arg \max_{k \in \{0, \dots, N\}} \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))] \right).$$

For a market of size  $N$ ,  $k^*(N)$  describes woman  $w$ 's optimal point of truncation, given that the other women submit their true preference lists. If there are multiple optima, we conservatively select that which involves the least truncation. We now have the following theorem.

**Theorem 3.1.** *Let woman  $w$  have uniform beliefs and preferences linear in rank (or any strictly increasing, convex transformation of such preferences). Then  $\lim_{N \rightarrow \infty} \frac{k^*(N)}{N} = 0$ .*

Theorem 4.1 states that as the market size grows large, the fraction of the list that an individual woman optimally truncates goes to 100%.<sup>15</sup> Note that under uniform beliefs, Roth and Rothblum's optimality theorem applies. Hence, the aggressive truncation strategies described in the theorem are the best overall strategies, not just the optimal truncation strategies.

The intuition behind this theorem can be gleaned from statistical facts about the most and least preferred achievable mates for women. In large markets where preferences are uniform, the expected rank of the most preferred achievable mate of a woman (which is the same as the expected rank

---

<sup>15</sup>We thank a referee for pointing out that this theorem will continue to hold if all agents drew uniformly at random preferences of length  $\alpha N$  for  $\alpha \in (0, 1)$ . We omit formal proof of this statement which would require slight adaptations to the theorems that we cite in our own proof.

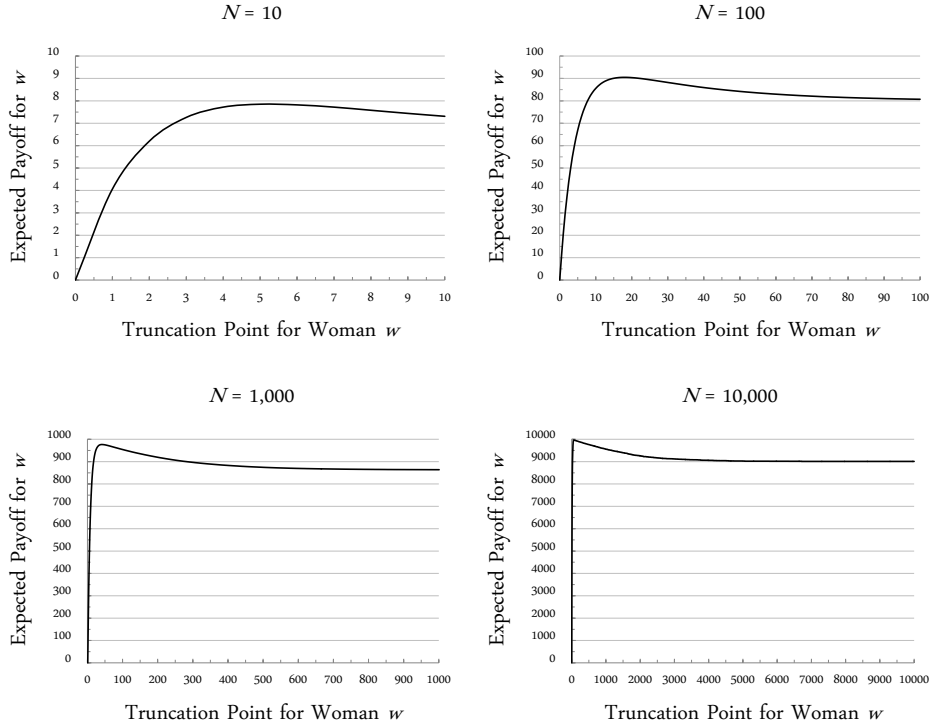
of her mate under *WP-DA*) is very low relative to the length of her list; it asymptotes to  $\log N$  (Pittel, 1989). This suggests that a woman may safely truncate a large fraction of her list with little risk of becoming unmatched. Furthermore, the expected rank of a woman's match under *MP-DA* is significantly worse, asymptoting to  $\frac{N}{\log N}$  (Pittel, 1989). In fact, for large markets, Pittel (1992) proved that the worst-off wife will be matched with a husband at the bottom of her list with probability approaching 1. This large gap in a woman's expected most and least preferred achievable mates suggests that not only is it safe to truncate a large fraction of one's list in large markets, but that a woman will also generate gains from such a truncation. See the appendix for the details of the proof, and Chapter 4 for a discussion of the unbalanced case in which these properties cease to hold (Ashlagi *et al.*, 2013).

To get a sense of the impact of truncation and to see examples of optimal truncation levels, we simulate markets of size  $N = 10, 100, 1,000$ , and  $10,000$  (Figure 3.1). In each market, we randomly generate a full preference list for each agent, and calculate an individual woman's payoff from truncating at each point in  $\{0, \dots, N\}$ , where a woman's payoff is given by  $(N + 1) -$  her partner's rank. We then iterate 100,000 times and average her payoffs.<sup>16</sup>

Observe that under truthful reporting,  $w$ 's payoff (given by the right hand side intercepts) is very close to  $\frac{N}{\log N}$ , the asymptotic limit found in Pittel (1989). Even in the largest market we simulated,  $N = 10,000$ ,  $w$ 's payoff at

---

<sup>16</sup>Formally, we are estimating  $\mathbb{E}[v(k, \tilde{P}_{-w})]$ , the expectation of a random variable, by averaging many independent draws of  $v(k, \tilde{P}_{-w})$ . With 100,000 draws, the 95% confidence intervals around each estimate of  $\mathbb{E}[v(k, \tilde{P}_{-w})]$  are so small that they are imperceptible when drawn on the graphs. For example, when  $N = 1,000$  and  $k = 500$ , the estimate of  $\mathbb{E}[v(500, \tilde{P}_{-w})]$  is 874.5, and the 95% confidence interval is (873.8, 875.2).



**Figure 3.1: *Simulation Results for Truncation Payoffs.*** The graphs display  $(N + 1)$ — an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, N\}$  and submitting these preferences to MP-DA. Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. Markets are of size 10, 100, 1,000 and 10,000.

the peak is roughly 10% higher than her payoff from truthful reporting. Note further that in each market, peak utility is lower than  $N + 1 - \log N$ , the asymptotic expected rank of a woman’s most preferred achievable mate. A woman will never be able to do better than this, even with perfect information about reported preferences of others.

In each of the graphs in Figure 3.1, and especially when  $N = 1,000$  and  $N = 10,000$ , there is a flat area on the right hand side. These lower levels of

truncation are unlikely to have any impact on payoffs – indeed establishing this is a crucial step in the proof of Theorem 4.1. Additional truncation can generate better mates, but still bears little risk of over-truncation. Finally, for extreme levels of truncation there is a high probability of over-truncation, leading to a steep dropoff in payoffs. As  $N$  grows larger, the “safe range” increases: we obtain larger flat zones and peaks moving to the left.

### 3.4 Truncation Equilibria and Welfare

In this section we consider the Bayesian game in which agents must submit preference lists to *MP-DA*. We demonstrate that in *equilibria in truncation strategies*, compared to outcomes from truthful preference list reporting, welfare for men is lower, welfare for women is greater, and the expected number of matches is lower. When there are multiple equilibria that can be compared in degree of truncation, women prefer the equilibrium where they truncate most, while men prefer the equilibrium where they truncate least. In *uniform markets*, we demonstrate the existence of a symmetric equilibrium in truncation strategies, but asymmetric equilibria may also exist. In a truncation equilibrium where some women truncate more than others, the women who truncate less receive higher payoffs. That is, while across equilibria women prefer to see higher degrees of truncation, within an equilibrium, they prefer not to be the ones bearing the risk of truncating.



### 3.4.1 Equilibria under General Preferences

Define the *Preference List Submission Game* as follows: Let  $U$  be a finite subset of the set of all possible utility profiles for  $I = \mathcal{M} \cup \mathcal{W}$  and  $\phi(\cdot)$  be a distribution over  $U$ . Let the message space of any agent  $i \in I$  be  $\mathcal{P}_i$ , the set of all possible strict preference lists for player  $i$ , with  $\mathcal{P} = \prod_i \mathcal{P}_i$ . Recall that  $\mu^M[P]$  gives the *MP-DA* matching for reported preference lists  $P$ . The *Preference List Submission Game* is the Bayesian game described by

$$\langle I, \mathcal{P}, \mu^M[\cdot], U, \phi(\cdot) \rangle.$$

A pure strategy for agent  $i$  is a mapping  $s_i : U_i \rightarrow \mathcal{P}_i$ , and a mixed strategy for  $i$  is a mapping  $\sigma_i : U_i \rightarrow \Delta(\mathcal{P}_i)$  which describes a randomization over submitted preference lists for each possible type. Define a *truncation strategy* for woman  $w$  as a strategy in which  $w$  mixes over truncations of her true preference list. For any two truncation strategies  $\sigma_w$  and  $\sigma'_w$  for a woman  $w$ , we say that  $\sigma_w$  involves more truncation than  $\sigma'_w$  if the distribution over truncation points induced by  $\sigma_w$  first order stochastically dominates the distribution induced by  $\sigma'_w$ .

We will restrict attention to equilibria where men report preferences truthfully, an assumption motivated by the dominant-strategy result of Dubins and Freedman (1981) and Roth (1982).<sup>17</sup> Define a Bayesian Nash equilibrium  $\sigma = (\sigma_{m_1}, \dots, \sigma_{m_N}, \sigma_{w_1}, \dots, \sigma_{w_N})$  in which men report truthfully and women mix over truncation strategies as an *equilibrium in truncation strategies*. The following theorem describes welfare in such equilibria.

---

<sup>17</sup>If we ignore this requirement, there is always a trivial equilibrium in which all players submit an empty list.

**Theorem 3.2.** *Let  $\sigma$  and  $\sigma'$  be equilibria in truncation strategies in which each woman truncates more under  $\sigma$  than under  $\sigma'$ . Then compared to the outcomes in  $\sigma'$ , under  $\sigma$ ,*

- i) welfare for women is weakly greater*
- ii) welfare for men is weakly lower*
- iii) the expected number of matches is weakly lower.*

*Furthermore, under both  $\sigma$  and  $\sigma'$ , i), ii) and iii) hold in comparison to the outcomes from truthful reporting of preferences to MP-DA.*

The results of Theorem 3.2 can be obtained by considering the effect of incremental truncations in light of Proposition 3.1. An incremental truncation by a woman  $w$  can only negatively affect the welfare of men: a “divorced” man will receive a worse mate, and the chain of offers that follows can only lead to worse mates for the other men as well. Since the chain will end in an offer accepted by some woman, or else in no match, the incremental truncation weakly decreases the number of matches. This logic underpins results *ii)* and *iii)*. At the same time, incremental truncation by a woman has a (weakly) positive spillover on the welfare of other women: rejection of a man can only lead to more offers for other women. The spillover from the truncation of other women, together with her best response requirement, imply that each woman weakly prefers the equilibrium with more truncation, and that any truncation equilibrium is preferred to truthful reporting.

Theorem 3.2 is similar in spirit to Kojima (2006) and Konishi and Ünver (2006) who show that in games of capacity manipulation in hospital-intern markets, every hospital prefers a Nash equilibrium to any reported profile

of larger capacities. Theorem 3.2 also brings to mind the welfare result in Coles *et al.* (2013), in which signaling equilibria with varying cutoffs are compared. In each of these settings, actions by one side of the market – capacity reduction in Kojima (2006) and Konishi and Ünver (2006), signaling by women in Coles *et al.* (2013), and truncation in this paper – serve to “shift the balance of power.” When there are equilibria with varying degrees of action, the sides of the market are at odds over which equilibrium is preferred, and whether any action is desirable at all.<sup>18</sup>

### 3.4.2 Equilibria in Uniform Markets

Let a *uniform market* be the setting in which each agent is equally likely to have any full preference list. Additionally, agent utility depends on partner rank, agents identically value a match with their  $r$ th ranked choice  $\forall r \in \{1, \dots, N\}$ , and have identical value to being unmatched.

**Theorem 3.3.** *In uniform markets, there exists a symmetric equilibrium  $((\sigma_m)^N, (\sigma_w)^N)$  where men each use the strategy  $\sigma_m$  of truthful reporting and women each use the strategy  $\sigma_w$ , which is a mixture over truncation strategies.*

**Proof.** We begin by constructing an auxiliary game. In this game, the set of players is the same as in the original game, the set of pure strategies for each woman is  $\{0, 1, \dots, N\}$ , and men all have one strategy,  $\{N\}$ . States of the world are profiles of preferences, which are realized with the same

---

<sup>18</sup>Another paper that bears mention is Ashlagi and Klijn (2012), which considers “group manipulations in truncation strategies” by women in the *MP-DA*. Such manipulations weakly benefit other women and harm other men. The results in Ashlagi and Klijn (2012) differ from ours, as we focus on equilibria and on incomplete information.

probability distribution as in the original game, but now players learn neither the preferences of others, nor their own preferences. Payoffs are defined according to the same utility function as in the original game, where each player receives the payoff from being matched to his stable partner under the profiles truncated at levels corresponding to the pure strategies chosen.

A standard argument due to Nash (1951) shows that the auxiliary game has an equilibrium, symmetric with respect to women. It is easy to see that this remains an equilibrium in the game where players observe their own preferences (but not the preferences of others) before choosing an action (truncation). Finally, returning to the unrestricted game, we recall Roth and Rothblum's optimality of truncation theorem from Section 4.2.3. Since men are playing dominant strategies, and since the strategies yield  $\mathcal{M}$ -symmetric beliefs, we conclude that the profile of strategies that we found is an equilibrium in the game where strategies are unrestricted. If it were not, then some woman could do strictly better by using a non-truncation strategy. But since each woman  $w$ 's beliefs in this setting are  $\mathcal{M}$ -symmetric, a truncation strategy weakly dominates this non-mixed strategy, which yields a contradiction.<sup>19</sup>  $\square$

In addition to symmetric equilibria, asymmetric equilibria may exist. The following example illustrates this.

---

<sup>19</sup>Using the same proof technique and following Ehlers (2008), we can also prove that in uniform matching markets that use an anonymous mechanism satisfying positive association, individual rationality, and independence of truncations, there exists a non-trivial equilibrium in which all men play the same truncation strategy and all women play the same truncation strategy. This class of mechanisms includes all priority mechanisms and all linear programming mechanisms introduced in British entry-level medical markets and in public school choice in some American cities.

**Example 3.2.** Consider a  $2 \times 2$  uniform market. Suppose each woman derives utility 10 from being matched to her top choice, 1 from being matched to her second choice, and 0 otherwise. We first calculate the probability of three events:

$A = \{\text{There is a unique stable matching, and in this matching } w_1 \text{ is matched to her top choice}\}$

$B = \{\text{There is a unique stable matching, and in this matching } w_1 \text{ is matched to her second choice}\}$

$C = (A \cup B)^c = \{\text{There are two stable matchings}\}$

A simple calculation shows that  $P(A) = \frac{5}{8}$ ,  $P(B) = \frac{2}{8}$  and  $P(C) = \frac{1}{8}$ .

Now suppose agents report preferences to *MP-DA*. If the other agents are truthful,  $w_1$  should truncate her list to include only her most preferred man (thus earning  $\frac{6}{8} \times 10 > \frac{5}{8} \times 10 + \frac{3}{8} \times 1$ ). But if  $w_1$  truncates her list in this manner,  $w_2$  has no incentive to truncate at all. Even if it turns out there was room for beneficial truncation (event  $C$ ),  $w_1$  has already done the “hard work” of truncating. She bears the risk of becoming unmatched, but also shifts the outcome from one matching to the other, improving payoffs for both women.  $\square$

Several observations can be made from this example. First, payoffs are higher in this equilibrium than under truthful reporting, as predicted by Theorem 3.2. Second, when  $w_1$  truncates more,  $w_2$  prefers to truncate less. While we don’t have a result that truncation under uniform preferences is a case of strategic substitutes, this example (and simulation evidence in Figure 3.3) stand in contrast to the complete information world where truncation

strategies are strategic complements. There, when some woman  $w_i$  truncates, this can only improve (or leave unchanged) the most preferred achievable mate for each  $w_j$ . Under complete information, this translates to a (weakly) greater optimal degree of truncation.

Another observation from the example is that  $w_2$ 's utility is  $\frac{6}{8} \times 10 + \frac{2}{8} \times 1$ , which is greater than the utility of  $w_1$ , which is  $\frac{6}{8} \times 10$ . That is, the agent who truncates less has greater utility than the agent who truncates more. In uniform markets, this result generalizes: within asymmetric truncation equilibria, we have a crisp preference among women against truncation. Theorem 3.4 encapsulates this.

**Theorem 3.4.** *Consider any asymmetric equilibrium in a uniform market where  $w_1$  truncates more than  $w_2$  (in the sense of first order stochastic dominance). Then i) if  $w_1$  and  $w_2$  swap strategies, the resulting profile will also be an equilibrium and ii)  $w_2$  prefers the original equilibrium, in which she truncates less.*

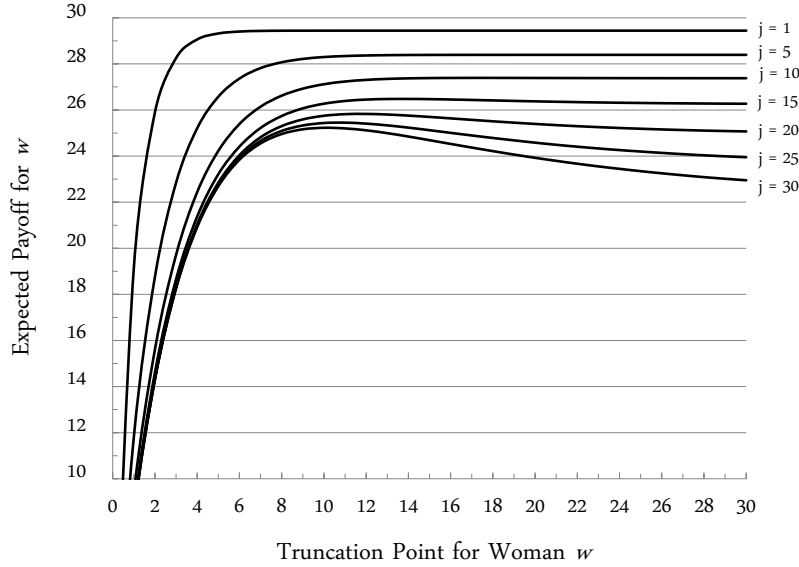
Intuitively,  $w_1$  and  $w_2$  face the same opposition except for one feature: each woman “competes” with the other, but not with herself. Woman 2, who truncates less, benefits from facing competition in which the other woman truncates more. Given that  $w_1$  is willing to take the risk of this truncation,  $w_2$  no longer feels compelled to do so herself.<sup>20</sup>

---

<sup>20</sup>A more general version of Theorem 3.4 also holds. For any Preference Submission Game with two “symmetric” women, i) and ii) remain valid.

### 3.4.3 Simulations: Finding a Symmetric Equilibrium

In this section we run simulations to explore equilibria in a uniform market. We assume that agents care about the rank of their partners (as in Figure 3.1, graphed as  $N + 1$ -rank, so that the graphs display maxima rather than minima), and we examine a market with  $N = 30$ . The simulations suggest that under these assumptions, there is a pure strategy symmetric equilibrium with a common truncation point that involves a non-trivial degree of truncation.



**Figure 3.2:** *Woman  $w$ 's expected payoff in a uniform market as a function of her truncation point when women  $\mathcal{W} \setminus \{w\}$  truncate at  $j$ .  $N = 30$ . Iterations = 1,000,000.*

We first examine how returns to truncation for  $w$  change when other women also truncate their lists. In Figure 3.2 we examine the effect on  $w$ 's payoffs when women  $\mathcal{W} \setminus \{w\}$  all truncate at a common point  $j$ , where  $j$  takes

on various values. For lower  $j$ , curves for  $w$  are higher. This follows from the positive spillover of truncation: when  $\mathcal{W} \setminus \{w\}$  truncate their lists, this benefits woman  $w$ .

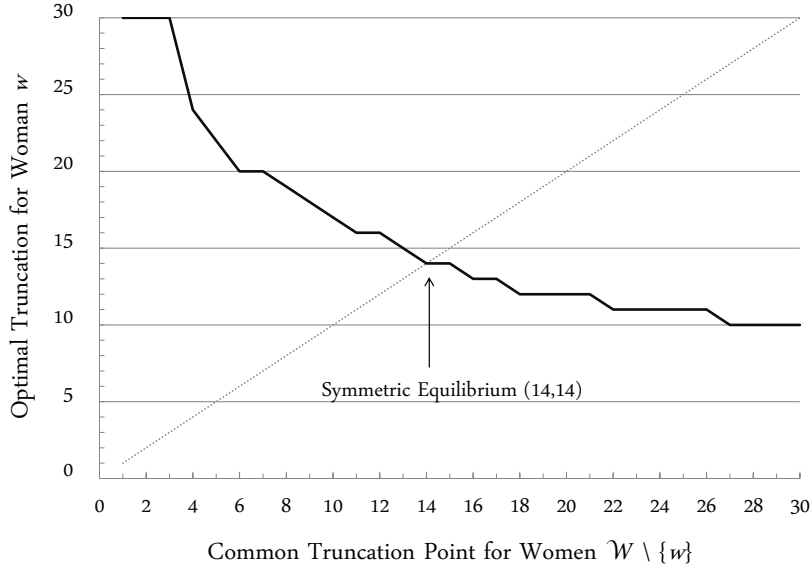
In Figure 3.2 it is also apparent that as  $w$ 's opponents truncate more, e.g., from  $j = 30$  to  $j = 25$  to  $j = 15$ , etc.,  $w$  should truncate less: as  $j$  increases, peaks move to the right. As in Example 3.2, this stands in contrast to the complete information result where truncation strategies are strategic complements. Note that at the extreme, when  $j = 1$ ,  $w$  can never benefit from truncation. Since truncation still bears risk, her optimal degree of truncation in this case is  $N = 30$ .

Observe also that when  $\mathcal{W} \setminus \{w\}$  submit very short lists, e.g.,  $j \in \{1, 5, 10\}$ ,  $w$ 's optimal truncation point is hard to observe because her payoff curve becomes very flat. The reasons for this flatness are two-fold: When other women submit very short lists, the expected number of stable matchings is known to be small (see Immorlica and Mahdian, 2005). Hence, there are minimal opportunities for beneficial truncation. At the same time, when other women truncate, the expected partner rank for  $w$  is very low (favorable). This leaves little danger that moderate levels of truncation will leave  $w$  unmatched. The minimal rewards and risks to truncation lead to the flatness of payoff curves.

By running a very large number of iterations, we identify the peaks of the curves in Figure 3.2. This exercise corroborates the hypothesis that under uniform preferences, truncation is a case of strategic substitutes. As illustrated in Figure 3.3, the optimal truncation point for  $w$  is inversely related to  $j$ , the common truncation point of women  $\mathcal{W} \setminus \{w\}$ . Of course, due



to the flatness of the expected payoff graph, optimal truncation points for small  $j$  are “just barely” optimal.



**Figure 3.3:** *Optimal truncation point of woman  $w$  in a uniform market as a function of the common truncation point  $j$  of women  $\mathcal{W} \setminus \{w\}$ .  $N = 30$ . Iterations = 10,000,000.*

By overlaying the 45° line, we identify the point of truncation in a pure strategy symmetric equilibrium to be 14, more than a 50 percent truncation of the entire list.<sup>21</sup> When all women truncate at this common point, no single woman sees significant gains from truncation compared to truthful reporting. However, since truncation has a positive externality on other women, the

<sup>21</sup>To test whether this is indeed a symmetric equilibrium, we “guess and verify.” We repeatedly sample  $w$ ’s payoffs under  $(k, j)$  where  $k$  is  $w$ ’s truncation point and  $j$  is the common truncation point of  $\mathcal{W} \setminus \{w\}$ , and establish that  $w$ ’s payoffs under  $(14, 14)$  are sufficiently distant from those under  $(15, 14)$ ,  $(13, 14)$ , and other profiles. To do this, we construct confidence intervals around  $(14, 14)$  and, using different draws, around other profiles. We observe that these intervals have an empty intersection.

equilibrium payoff is non-trivially greater than the payoff should all women report truthfully (see Figure 3.2).

This equilibrium potentially leads to ex-post instability with respect to true preferences. Some women might over-truncate and be left unmatched. More subtly, when *other* women truncate, woman  $w$  may be paired with a mate that is not achievable under the true preferences.

These instabilities suggest a possible application of these results: the impact of strategic behavior on a post-market “scramble” for positions. Since truncation can lead to unmatched participants following the match, a second, organized match might be helpful to find partners for these agents. Indeed, post-market scrambles have been organized in both the market for medical residents as well as in the job market for new economists.<sup>22</sup>

At first observation, an organized scramble would reduce the downside to remaining unmatched in the primary match. But this might induce additional risk-taking behavior (more truncation) by participants. Such behavior would increase the pool size in the second match, raising the value of being unmatched, inducing even more truncation. A secondary match might ultimately enjoy high participation levels, but only because it has drawn participants away from the primary match, complicating overall welfare analysis.

---

<sup>22</sup>The NRMP offers the “Supplemental Offer and Acceptance Program” (SOAP), which replaced a somewhat less orderly scramble (see <http://www.nrmp.org/2012springmeeting.pdf>). The American Economic Association organizes the “Scramble” in which candidates seeking jobs and employers with positions open late in the job market can announce their availability on the AEA website (see <http://www.aeaweb.org/joe/scramble>).

### 3.5 Truncation and Risk Aversion

While gains to truncation can be significant, truncation is nevertheless a risky strategy. When  $w$ 's opponents truncate, truncation for  $w$  is particularly risky: compared to truthful reporting, optimal truncation offers minimal benefit, and over-truncating can lead to large losses. One might expect agents with more conservative attitudes toward risk to shy away from this proposition. In this section, we ask how a woman's truncation behavior varies as we vary her attitude towards risk.

We consider a general setting, with arbitrary preferences for woman  $w$  and beliefs about reported preferences of others. Let  $\psi(\cdot)$  be any strictly increasing, concave transformation. We will show that for any beliefs about others, woman  $w$  with preferences  $u_w(\cdot)$  will truncate more than a woman  $w_\psi$  who has identical beliefs, but preferences given by  $\psi(u_w(\cdot))$ .

Recall that when we fix  $w$ 's preferences to be  $u_w(\cdot)$ , we defined shorthand

$$v(k, P_{-w}) \equiv u_w(\mu^M[P_w^k, P_{-w}](w)),$$

her payoff from submitting truncated preference list  $P_w^k$ . Now define

$$v_\psi(k, P_{-w}) \equiv \psi(u_w(\mu^M[P_w^k, P_{-w}](w))),$$

the payoff from submitting truncated preference list  $P_w^k$  for a woman  $w_\psi$  with preferences  $\psi(u_w(\cdot))$ .

The following theorem states that if  $w$  prefers truncating less to more, then  $w_\psi$  definitely prefers truncating less to more.

**Theorem 3.5.** *Let  $\tilde{P}_{-w}$  be any random variable distributed over  $\mathcal{P}_{-w}$ . Then*

$\forall k \in \{1, \dots, N-1\}, \forall t \in \{1, \dots, N-k\}$  we have

$$\begin{aligned} \mathbb{E}[v(k, \tilde{P}_{-w})] &\leq \mathbb{E}[v(k+t, \tilde{P}_{-w})] \Rightarrow \\ \mathbb{E}[v_\psi(k, \tilde{P}_{-w})] &\leq \mathbb{E}[v_\psi(k+t, \tilde{P}_{-w})]. \end{aligned}$$

Furthermore, if i)  $\psi(\cdot)$  is strictly concave, and ii) under  $\tilde{P}_{-w}$ , each man is achievable for  $w$  with positive probability, then the second inequality is strict.

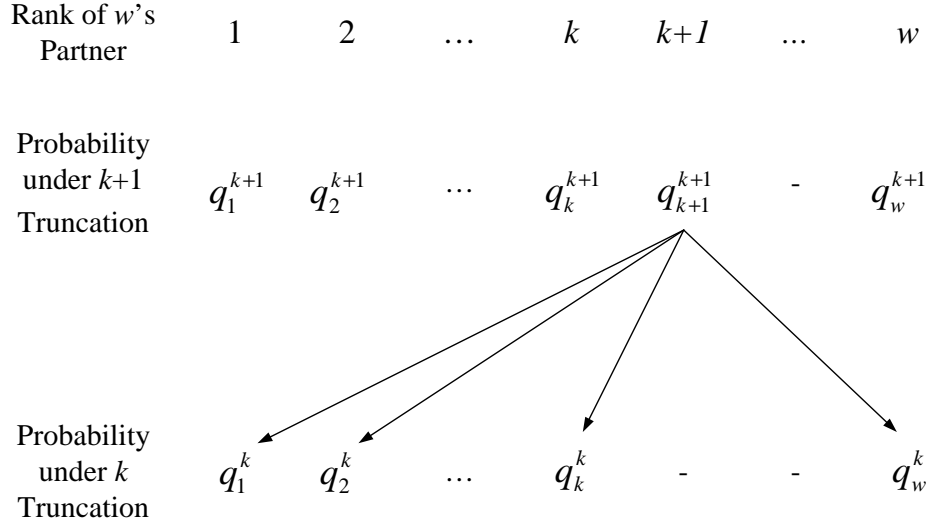
The constructive proof nicely illustrates incremental truncation analysis, so we provide it in-text.

**Proof.** We begin with the proof for  $t = 1$ . An analogous argument works for all other  $t$  in the given range, with necessary proof adjustments described at the end. Our technique focuses on two lotteries over outcomes. Let  $Q^{k+1}$  be the lottery over mates for  $w$  when she truncates at  $k+1$ , and let  $Q^k$  be the lottery when she truncates at  $k$ . Our goal is now to show that if  $Q^k$  is mean-decreasing as compared to  $Q^{k+1}$  from  $w$ 's perspective (in terms of her von Neumann-Morgenstern utility), then it will be mean-decreasing from  $w_\psi$ 's perspective as well.

Distributions  $Q^{k+1}$  and  $Q^k$  are shown in Figure 3.4. Recalling Proposition 3.1,  $k$ -truncating is equivalent to  $(k+1)$ -truncating followed by  $k$ -truncating. That is, lottery  $Q^k$  is equivalent to starting with lottery  $Q^{k+1}$ , then rolling the die again if  $w$  receives her  $(k+1)$  ranked choice. Hence,

$$q_i^k \geq q_i^{k+1} \quad \forall i \in \{1, \dots, k\} \cup \{w\}.$$

Let shorthand  $u_i(Q)$ ,  $i \in \{w, w_\psi\}$  describe  $i$ 's expected utility from lottery  $Q$ . Suppose first that  $u_w(Q^k) = u_w(Q^{k+1})$ , that is,  $\mathbb{E}[v(k, \tilde{P}_{-w})] =$



**Figure 3.4:**  $k$ -Truncation is equivalent to  $(k+1)$ -truncation followed by an extra gamble: divorcing man  $k+1$ .

$\mathbb{E}[v(k+1, \tilde{P}_{-w})]$ , so that from  $w$ 's perspective,  $Q^k$  is a *mean-preserving spread* of  $Q^{k+1}$ . Then by Jensen's inequality,  $u_{w_\psi}(Q^k) \leq u_{w_\psi}(Q^{k+1})$ . If  $\psi(\cdot)$  is strictly concave and  $Q^k \neq Q^{k+1}$  (which follows from *ii*), then  $u_{w_\psi}(Q^k) < u_{w_\psi}(Q^{k+1})$ .

Now suppose that  $u_w(Q^k) < u_w(Q^{k+1})$ , so that from  $w$ 's perspective,  $Q^k$  is *mean-decreasing* as compared to  $Q^{k+1}$ . We will now construct an intermediate lottery  $Q'$  such that

1.  $Q^{k+1}$   $P_w$ -stochastically dominates  $Q'$  and
2. From  $w$ 's perspective,  $Q^k$  is a mean preserving spread of  $Q'$ .

Define lottery  $Q'$  so that  $Q'$  is identical to  $Q^{k+1}$ , except that we replace outcome  $k+1$  ( $w$ 's  $(k+1)$  ranked choice) with lottery  $\alpha(k+1) + (1-\alpha)w$ . Choose  $\alpha \in [0, 1]$  such that  $w$  has  $u_w(Q') = u_w(Q^k)$ . Such an  $\alpha$  must exist:

when  $\alpha = 1$ ,  $Q' = Q^{k+1}$ , and when  $\alpha = 0$ ,  $u_w(Q') \leq u_w(Q^k)$ . Our desired  $\alpha$  follows from the Intermediate Value Theorem.

By construction,  $Q^{k+1}$   $P_w$ -stochastically dominates  $Q'$ . With respect to  $w$ 's utility, we also have that  $Q^k$  is second order stochastically dominated by  $Q'$ . To see this, observe that  $Q'$  was constructed to have the same mean as  $Q^k$ , and that compared to  $Q'$ ,  $Q^k$  shifts probability mass to the extremes:  $q_j^k \geq q_j'$  for  $j \in \{1, \dots, k\} \cup \{w\}$ .

Since  $P_w = P_{w_\psi}$ , by  $P_w$ -stochastic dominance,  $w_\psi$  also strictly prefers  $Q^{k+1}$  to  $Q'$ . By Jensen's inequality,  $w_\psi$  weakly prefers  $Q'$  to  $Q^k$ . Hence,

$$u_{w_\psi}(Q^k) < u_{w_\psi}(Q^{k+1}),$$

so the theorem is proved for  $t = 1$ .

When  $t > 1$ , we may again construct an intermediate lottery  $Q'$ , this time that transfers weight from  $\{k+1, k+2, \dots, k+t\}$  to the unmatched option  $w$ . Just as before, we can construct  $Q'$  to ensure that  $Q^{k+1}$   $P_w$ -stochastically dominates  $Q'$ , and that from  $w$ 's perspective,  $Q^k$  is a mean preserving spread of  $Q'$ . The key insight is that truncation transfers probability mass to the extremes: the most preferred mates, as well as the unmatched option.  $\square$

We can now use Theorem 4.5 to sort optimal truncation points based on degree of concavity.

**Corollary 3.1.** *Let  $k_i^l$  be the minimum optimal truncation point (by rank) and let  $k_i^h$  be the maximum optimal truncation point for woman  $i \in \{w, w_\psi\}$ . Then  $k_w^l \leq k_{w_\psi}^l$  and  $k_w^h \leq k_{w_\psi}^h$ . Furthermore, if conditions i) and ii) from Theorem 4.5 hold, then  $k_w^h \leq k_{w_\psi}^l$ .*

**Proof.** If  $k_w^l$  is  $w$ 's minimum optimal truncation point, then  $w$  strictly prefers truncation at  $k_w^l$  to truncation at any  $k < k_w^l$ . Following the reasoning of Theorem 4.5,  $w_\psi$  must then prefer truncation at  $k_w^l$  to truncation at any  $k < k_w^l$ . Hence,  $k_w^l \leq k_{w_\psi}^l$ . A similar argument can be used to show  $k_w^h \leq k_{w_\psi}^h$ .

If  $k_w^h$  is  $w$ 's maximum optimal truncation point, then  $w$  (weakly) prefers truncation at  $k_w^h$  to truncation at any  $k < k_w^h$ . If conditions *i*) and *ii*) hold, then  $w_\psi$  must strictly prefer truncation at  $k_w^h$  to truncation at any  $k < k_w^h$ . Hence,  $k_w^h \leq k_{w_\psi}^h$ .  $\square$

Thus, when facing the same environment, players who are more risk averse truncate less, with the set of optimal truncation points overlapping at the very most at one point.

The key insight in the analysis is the interpretation of truncation as a risky lottery, and then mapping the additional risk associated with incremental truncation to an *extra lottery* a woman must face. If a woman doesn't like to face the extra lottery, then certainly a woman with more concave preferences will not want to face it.

Note that despite pertaining to risk aversion, the results in this section do not restrict the structure of  $u_w(\cdot)$  in any way. For example, we do not require  $u_w(\cdot)$  to be "concave." Rather, it is the *relative* concavity that is crucial. For example, if we restrict ourselves to the class of functions that are *s-shaped* in rank, we know that within this class, concave transformations induce less truncation.

In a general sense, this result can be taken as advice to participants. Players can observe the patterns of behavior of others, size up their own

attitudes toward risk, and truncate more or less accordingly. In markets where there is a steep dropoff in utility from a woman’s most preferred partner to her second choice, a smaller dropoff from choice two to three, and so forth, we may anticipate more aggressive truncation. On the other hand, if participants are largely content with any of the available choices, but see great disutility from being unmatched, truncation is not advisable.

This result can also offer advice to a market designer. If an objective is to maximize the number of matches, a market designer may wish to choose the less risk averse side to be the “proposers” in the Deferred Acceptance Algorithm. If the two sides of the market are identical in all regards except for their risk preferences, the more risk averse side will be less likely to truncate, even if manipulations increase their expected partner rank. Lower levels of truncation will increase the number of realized matches, and consequently, reduce the number of participants left unmatched. However, in making this choice, the market designer should take other market features into consideration as well, as we demonstrate in the next section.

### 3.6 Correlated Preferences

In the Preference List Submission Problem for Women, we now let woman  $w$  believe other women in the market have preferences similar to hers. We consider how woman  $w$  should vary her degree of truncation as the degree of similarity varies. We provide evidence, both theoretical and simulation-based, that the greater the similarity in the preferences of other women to her own, the less woman  $w$  should truncate.



### 3.6.1 Perfectly Correlated Preferences

We consider first the case of perfectly correlated preferences on the women's side of the market.

*Remark.* When women have identical preferences, there is a unique stable matching.

To see this, note that the top-ranked man, as agreed upon by all women, must be matched with his most preferred partner in any stable matching, or else these two would constitute a blocking pair. The second-ranked man must then be matched to his most preferred remaining woman, and so on. *MP-DA* reduces to a serial dictatorship, determined by the common ranking of the men.

Since there is a unique stable matching in this setting, an individual woman's misrepresentation of her preference list can never improve her match. In fact, if a woman is certain that other women share her preferences (and are reporting truthfully), but is uncertain about what men will submit to the algorithm, truncation can very well lower her outcome by leaving her unmatched.

### 3.6.2 Partially Correlated Preferences

In this section, we introduce a notion of partial correlation of preferences indexed by a single parameter  $\alpha$ . We will show that the greater the degree of correlation, the less a woman should truncate.

Consider the Preference List Submission Problem for woman  $w$  with preferences  $u_w$  and beliefs  $\tilde{P}_{-w}$  about reported preference lists of opponents.

Let  $p(\cdot, \cdot)$  be the probability mass function for  $w$ 's beliefs. That is,

$$p(P_{\mathcal{M}}, P_{\mathcal{W} \setminus \{w\}})$$

gives the likelihood that the men will report preference lists  $P_{\mathcal{M}}$  and women  $\mathcal{W} \setminus \{w\}$  will report preference lists  $P_{\mathcal{W} \setminus \{w\}}$ . Define the marginal probability over mens' preference profiles by  $p^M(\cdot)$ .

Given  $p(\cdot, \cdot)$ , define beliefs  $p^C(\cdot, \cdot)$  by

$$p^C(P_{\mathcal{M}}, P_{\mathcal{W} \setminus \{w\}}) \equiv \begin{cases} p^M(P_{\mathcal{M}}) & \text{if } P_{\hat{w}} = P_w \quad \forall \hat{w} \in \mathcal{W} \setminus \{w\} \\ 0 & \text{otherwise} \end{cases}.$$

$p^C(\cdot, \cdot)$  is the distribution that preserves the marginal distribution over men's preferences  $p^M(\cdot)$ , but where the other women share the preferences of  $w$ .

Define beliefs  $p^\alpha(\cdot)$  by

$$p^\alpha(P_{-w}) \equiv (1 - \alpha)p(P_{-w}) + \alpha p^C(P_{-w}).$$

Hence, as  $\alpha$  varies from 0 to 1,  $p^\alpha$  ranges from  $p$  to  $p^C$ . The marginal distribution over men's preferences remains fixed, while the correlation of women's preferences steadily increases (the distribution remains constant if  $p = p^C$ ).

The set of optimal truncation points for woman  $w$  with preferences  $u_w$  and beliefs indexed by  $\alpha$  is given by

$$k^*(\alpha, p, u_w) \equiv \arg \max_{k \in \{0, \dots, N\}} \mathbb{E}_{p^\alpha}[v(k, \tilde{P}_{-w})].$$

Notice that since the choice set is finite,  $k^*(\cdot, \cdot)$  will be non-empty.

Let  $k^h(\alpha, p, u_w) = \max[k^*(\alpha, p, u_w)]$  and  $k^l(\alpha, p, u_w) = \min[k^*(\alpha, p, u_w)]$ ,

the optimal choices involving the least and most truncation respectively.

The following proposition states that for any preferences  $u_w$  and beliefs  $p$ , as we increase the degree of correlation  $\alpha$ , woman  $w$  should truncate less.

**Proposition 3.2.** *Let  $\alpha, \alpha' \in [0, 1]$  with  $\alpha' > \alpha$ . Then  $k^l(\alpha', p, u_w) \geq k^l(\alpha, p, u_w)$  and  $k^h(\alpha', p, u_w) \geq k^h(\alpha, p, u_w)$ .*

The proof relies on the fact that when there is a unique stable matching, it can never hurt to submit a full list. Using this fact, we can show that if under low correlation,  $w$  prefers truncating less to more, then under high correlation  $w$  definitely prefers truncating less to more. This is enough to sort optimal truncation points.

Intuition for this result is related to the size of the set of stable matchings. Truncation can yield improvement only when there are multiple stable matchings. The greater the degree of correlation, the smaller this set, and the lower the likelihood that a window for gain from truncation exists.

The anticipated level of correlation in the environment might influence the advice a market designer offers participants. If correlation is high, the designer can safely advise participants to report truthfully, and it is in their best interest to do so. With low correlation (sufficiently heterogeneous preferences), players may anticipate gains from truncation, which if acted on, could lead to unstable matchings.

### 3.6.3 Noisy Preferences

In Section 3.6.2, a woman believes it is possible that opponents have preference lists identical to hers. In this section, woman  $w$  believes women have

preference lists similar to hers, but not necessarily identical. We model such beliefs for women by generating noisy deviations from a common preference list. By performing simulations, we corroborate the theoretical results in Section 3.6.2; more correlation means a woman should truncate less.

We generate correlated preferences as follows. Each man  $m_i$  is assigned a random number  $r_i \sim U[0, 1]$ , and this value is agreed upon by all women. For each man  $m_i$ , each woman  $w_j$  also assigns an idiosyncratic (noise) component,  $q_{ij} \sim U[0, 1]$ . Woman  $w_j$ 's rankings over men are then determined by the sum  $\alpha \cdot r_i + (1 - \alpha)q_{ij}$ , where  $\alpha \in [0, 1]$  is a parameter that we will vary. Observe that from the perspective of any woman  $w$ , the preferences of other women are noisy versions of her own rankings. Values of  $\alpha$  close to one imply low noise, so  $\alpha$  measures the degree of correlation. Men are assumed to have uniformly random rankings over the women.<sup>23</sup>

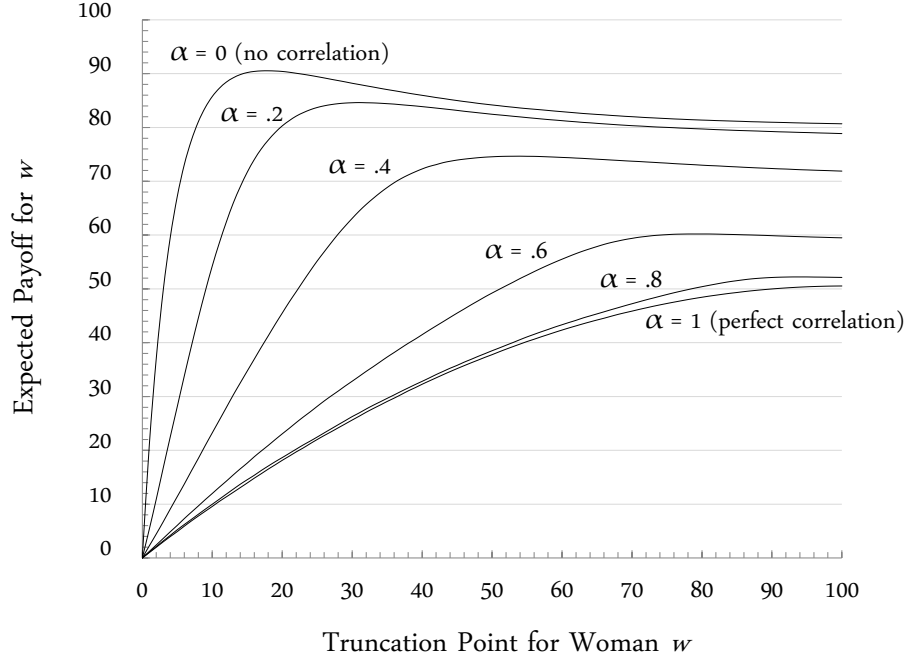
The process just described is used only to determine preference orderings. We further assume that  $w$ 's payoff is given by  $(N + 1 - \text{partner rank})$ , and being unmatched is just worse than being matched to her least preferred man, so we can compare outcomes to those depicted in Figure 3.1.

Figure 3.5 graphs the return to truncation for various values of  $\alpha$ . For each value, we randomly generated 100,000 preference list profiles and for each  $k$ , we graph woman  $w$ 's average payoff from  $k$ -truncation, when other agents are truthful.

When  $\alpha = 0$  (the top curve), this corresponds to uniform beliefs for  $w$ ,

---

<sup>23</sup>The common starting point for preferences might be an aggregate ranking based on available data, like the US News and World Report's annual ranking of universities. Caldarelli and Capocci (2000) simulate preferences in a one-to-one model similarly. In their model, the common component  $r_i$  is a man's "beauty," which in their view, evidently, is not in the eye of the beholder.



**Figure 3.5:** The graphs display  $(N + 1)$ — an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, N\}$  and submitting these preferences to MP-DA. Preference lists of men are uniformly random, and lists for women are randomly generated using the procedure described in the text. Payoffs are averaged over 100,000 draws.

the case studied in Section 3.3.4. When  $\alpha = 1$ , all women rank men the same way, the stable matching will be unique, and truncation cannot be helpful (as in Section 3.6.1).

From Figure 3.5, we make two key observations. First, woman  $w$  dislikes correlation. This fact is easy to explain. If all women agree on who the top men are, they “compete” for them as mates. The lower the correlation, the less the competition, and the better the expected mate for  $w$ . Second,  $w$ ’s optimal truncation point increases as correlation increases. This corroborates

the result in Section 3.6.2: when there is more correlation,  $w$  should truncate less.

### 3.7 Discussion and Conclusion

In this paper, we study optimal strategic behavior in one-to-one matching markets that are based on the Deferred Acceptance Algorithm, when agents have incomplete information about the preferences of others. We focus on truncation strategies. Among classes of strategies for preference list misrepresentation, truncation is an attractive option because it is guaranteed to weakly increase the likelihood of matching with one's more preferred partners. By contrast, more complicated strategies, such as swapping the order of agents in a preference list, may require detailed information about the preferences reported by others, and their outcomes are more difficult to predict.

Recent work by Immorlica and Mahdian (2005), Kojima and Pathak (2009) and others demonstrate that in large markets where agents submit short preference lists, opportunities for manipulation are limited. Lee (2011) presents a random utility model and shows that, in some sense, gains from manipulation become small in large markets.<sup>24</sup> In light of these findings one may ask whether agents – especially agents with little detailed information – can ever substantially gain from manipulation. Our paper answers in the affirmative. When agents view reported preference lists of others as being

---

<sup>24</sup>Lee's model is more general than ours, in many respects. Note, however that his model requires agent utilities to be bounded, while our model does not exclude unbounded functions.

drawn uniformly from the set of all possible full length preference lists, they may truncate their lists with little risk of being unmatched, but with the potential to see large gains in terms of the expected partner rank. Importantly, we show that while according to Lee (2011), utility gain from manipulation may be small, the optimal truncation may still be substantial. This finding provides an essential qualification to his results.

For many of the settings in which the Deferred Acceptance Algorithm has been successfully applied, notably in the NRMP and in the Boston and New York school systems, the markets do reflect large numbers and short preference lists. But the high levels of optimal truncation demonstrated in this paper raise a key issue: in large markets where agents submit short preference lists, can we be sure that the short lists were not simply the result of optimization? Costliness of information discovery often places natural limits on the length of submitted preference lists. Flyouts are costly for medical students; perhaps somewhat less so for hospitals. Nevertheless, this paper illustrates the theoretical possibility that even with full information about one's own preferences, substantial truncation (submission of short lists) may simply be utility-maximizing strategic behavior.

## Chapter 4

# Strategic Behavior in Unbalanced Matching Markets<sup>1</sup>

### 4.1 Introduction

A great success story in economic theory is the application of the *Deferred Acceptance Algorithm* (DAA), proposed by Gale and Shapley (1962), to real world two-sided matching markets. The DAA and its variants have been used extensively in school choice settings (Abdulkadiroğlu *et al.*, 2005), and most famously in the National Resident Matching Program (Roth and Peranson, 1999). The advantages of mechanisms using DAA over other mechanisms have been discussed extensively (See for example (Roth, 1990)). Importantly,

---

<sup>1</sup>Co-authored with Peter Coles



it was shown that while no stable matching mechanism is strategy proof, mechanisms applying the DAA have truthful reporting as a dominant strategy for the proposing side (Dubins and Freedman, 1981; Roth, 1982). The choice of the proposing side has received some attention in the public domain and in the literature (Roth and Peranson, 1999), but the general message that has emerged from this body of literature is that the choice of the proposing side has a small effect over agents' utilities (Roth and Peranson, 1999; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Lee, 2011; Ashlagi *et al.*, 2013; Azevedo and Leshno, 2012).

This paper takes a different perspective on this issue. We look for the (exact) best responses of agents, and consider the degree of manipulation expected in the market. To do this, we restrict attention to truncation strategies, which are endowed with a natural metric for measuring the extent of manipulation (how many acceptable partners were declared unacceptable). This class of strategies was shown to be optimal in symmetric low-information settings (Roth and Rothblum, 1999; Ehlers, 2004). We derive comparative statics on the extent of manipulation as a function of risk aversion and correlation, and show that more risk averse agents submit longer lists (so they are “more truthful”) and that correlation in preferences also reduces the incentives to manipulate. These results are similar to the findings of Coles and Shorrer (2014), but they are more general as we do not assume that the markets are balanced.

The main innovation in this paper is inspired by the results of Ashlagi *et al.* (2013). In contrast to the findings of Roth and Peranson (1999) regarding the “large core” of markets when agents have long preference lists, and the related

findings of Pittel (1989), Ashlagi *et al.* (2013) show that if the number of agents on each side of the market is not balanced, the core becomes small in the typical case. So while the gap between men (women) expected partner ranks under the men and women proposing versions of the DAA is high in a balanced market, even a slight imbalance “shrinks” this gap significantly. In light of this finding we ask: What are the effects of imbalance on the incentives to misrepresent one’s preferences? The answer is: it depends! Under the men-proposing version of the DAA, if there are more women than men, women optimally submit “long” lists. When the sides of the market are balanced, a woman facing truthful opponents should submit a short list; asymptotically she truncates 100% of her list. When women are over demanded (on the short side), we provide simulation evidence that extreme truncation is still optimal. We also show that truncation is “safe” when women are on the short side, but not when they are on the long side of the market. To summarize, the extent of optimal truncation may crucially depend on whether the strategic agents (the ones not on the proposing side) are on the long side or the short side of the market.

A market designer may prefer that agents submit either long or short lists. She may be concerned about the incentives for truthfulness for several reasons. For one, she may wish to advise participants that being truthful will not harm them, so as to “level the playing field” between savvy and naive agents (Pathak and Sonmez, 2008; Featherstone and Mayefsky, 2010). The number of matched agents and the (ex-post) stability of the match may also be affected (Featherstone and Niederle, 2011). An additional reason why designers may want to induce truthful reporting is that the submitted profiles

may provide a signal as to the desirability of agents on the two sides of the market. In school choice settings, for example, truthful reporting allows school districts to learn about the *actual* desirability of different schools (??).

But market designers may also have reasons to favor shorter lists. From a computational perspective, running the DAA on shorter lists is faster. More importantly, the designer may think that there is a cost (actual or mental) to the system or to participants on one side of the market for generating a long preference list. For example, a school may be required to give each applicant a tour, paperwork may be required for each school that appears on an applicant’s list, and a student may simply find it hard to compare his 100th and 101st choices. We take no stand on whether ensuring truthfulness or promoting short lists is more desirable, but merely wish to provide advice to the market designer given objectives regarding list length.

## 4.2 Preliminaries

We begin by setting out the basic model of matching. Following Coles and Shorrer (2014), and in contrast to some of the well-known papers in the field of matching, we endow agents with cardinal rather than ordinal preferences.

### 4.2.1 Marriage Markets and Stability

In this paper, only *one-to-one two sided matching markets* will be considered. We call these markets *marriage markets* for short, and label one side on the market as *men*  $\mathcal{M}$ , and the other as *women*  $\mathcal{W}$ . Both men and women are referred to as *agents*.

The *preferences* of man  $m \in \mathcal{M}$  are given by a von Neumann-Morgenstern utility function  $u_m : \mathcal{W} \cup \{m\} \rightarrow \mathbb{R}$ .  $u_m(w)$  is the utility that man  $m$  derives from being matched with woman  $w$  with certainty, and  $u_m(m)$  is his utility from being unmatched. For simplicity, we assume that  $u_m$  is one-to-one, so that there are no indifferences. Preferences for women are defined similarly. We denote by  $u = \prod_{i \in \mathcal{M} \cup \mathcal{W}} u_i$  the profile of agents preferences.

Since we have assumed that agents' preferences are one-to-one, they induce strict preference orderings on all possible partners and the possibility of remaining unmatched. For a man  $m \in \mathcal{M}$  we denote by  $P_m$  the preference list over  $\mathcal{W} \cup \{m\}$  that is induced by  $u_m$ . For example,  $P_m$  ranks  $w_3$  higher than  $w_1$  if  $u_m(w_3) > u_m(w_1)$ . We say that  $w \in \mathcal{W}$  is *acceptable* for  $m$  if  $u_m(w) > u_m(m)$ , so  $m$  prefers being matched with  $w$  over remaining single. We sometimes omit unacceptable mates from  $m$ 's preference list for notational convenience. Preference lists for women are defined similarly, and we denote by  $P$  the profile of all preference lists.

A matching  $\mu$  is a mapping from  $\mathcal{M} \cup \mathcal{W}$  to itself, such that for each  $m \in \mathcal{M}$  we have that  $\mu(m) \in \mathcal{W} \cup \{m\}$ , for each  $w \in \mathcal{W}$  we have  $\mu(w) \in \mathcal{M} \cup \{w\}$  and for each  $x \in \mathcal{M} \cup \mathcal{W}$   $\mu^2(x) = x$ . When  $\mu(x) = x$  we say that  $x$  is *single* or *unmatched* under the matching  $\mu$ . Otherwise, we refer to  $\mu(w)$  as  $w$ 's *husband* and  $\mu(m)$  as  $m$ 's *wife* under the matching  $\mu$ . We also use the terms *partner* and *mate*. The preferences over partners induce natural preference order over matchings, where each agent ranks the matchings according to the partner that is assigned to him.

A matching is *individually rational* if for every  $x \in \mathcal{M} \cup \mathcal{W}$ , the agent  $x$  weakly prefers  $\mu(x)$  to remaining single. A matching is *blocked* by a pair

$(w, m) \in \mathcal{W} \times \mathcal{M}$  if both  $w$  prefers  $m$  to  $\mu(w)$  and  $m$  prefers  $w$  to  $\mu(m)$ . A matching is stable if it is individually rational and not blocked by any pair. There always exists a stable matching in a market, but in general there may be more than one (Gale and Shapley, 1962). For given preferences, we say that a woman  $w$  is *achievable* for a man  $m$  if there exists a stable matching  $\mu$  such that  $\mu(w) = m$ . A symmetric definition applies to women's achievable mates.

#### 4.2.2 The Men-Proposing Deferred Acceptance Algorithm

To prove that every marriage market has a stable matching, Gale and Shapley (1962) proposed the Men-Proposing Deferred Acceptance Algorithm (*MP-DA*). It takes as an input a profile of preferences  $P$  of a set of agents  $\mathcal{M} \cup \mathcal{W}$  and outputs a stable matching  $\mu^M[P]$ . When  $P$  is clear from the contexts, we sometimes omit it and write  $\mu^M$  instead of  $\mu^M[P]$ . The following is a description of the algorithm.

- **Step 1.** Each man proposes to the first woman on his preference list. Each woman then considers her offers, rejects all men deemed unacceptable, and if any others remain, rejects all but her most preferred mate.
- **Step  $k$ .** Each man who was rejected in step  $k - 1$  makes an offer to the next woman on his preference list. If his preference list is exhausted, or if he prefers bachelorhood to the next woman on his list, he makes no offer. Each woman behaves as in step 1, considering offers in hand

(including any man she has retained from the previous step) and rejects all but her most preferred acceptable suitor.

- **Termination.** If in any step  $k$ , no man makes an offer, the algorithm terminates. Each woman is paired with her current mate and this matching is final.

Gale and Shapley show that this algorithm must terminate in finite time, and they provide a remarkable characteristic of the resulting outcome.

**Theorem.** (*Gale-Shapley*) *The matching  $\mu^M$  resulting from MP-DA is stable. Furthermore, for any other stable matching  $\mu$ , every man weakly prefers  $\mu^M$  to  $\mu$  and every woman weakly prefers  $\mu$  to  $\mu^M$ .*

Since there is no actual content to gender (it is just a label), it is clear that the women-proposing version of the algorithm (*WP-DA*) has identical but reversed properties. We denote its output given an input  $P$  by  $\mu^W[P]$ .

As discussed by Roth (1990), stability is a desirable property for a matching mechanism. But the theorem illustrates a particular feature of the stable matching produced by the MP-DA (WP-DA); it is the most desirable stable matching for men (women), and the least desirable for women (men). This paper focuses on the strategic incentives that emerge from this property under incomplete information, and their effects on the realized matchings given strategic reporting.

### 4.2.3 The Preference List Submission Problem

We now turn to study the incentive properties of stable matching mechanism which use the MP-DA. In a setting where agents are asked to report preferences lists to the mechanism, we consider if they have an incentive to report truthfully, or to submit a different preference list.

Consider a set of agents  $\mathcal{M} \cup \mathcal{W}$ . Agent  $i \in \mathcal{M} \cup \mathcal{W}$  with preferences  $u_i$  must submit a preference list  $\hat{P}_i$  to *MP-DA*, where  $\hat{P}_i$  is chosen from the set of  $i$ 's possible preference lists  $\mathcal{P}_i$ . The agent's beliefs about what preference lists others will report are represented by the random variable  $\tilde{P}_{-i}$ , which takes as its range  $\mathcal{P}_{-i}$ , the set of all possible preference list profiles for others. Note that since  $u_i$  is a von Neumann-Morgenstern utility function, agent  $i$  may compare outcomes in this incomplete information setting.

Agent  $i$  solves the *Preference List Submission Problem*:

$$\max_{\hat{P}_i \in \mathcal{P}_i} \mathbb{E}[u_i(\mu^M[\hat{P}_i, \tilde{P}_{-i}](i))].$$

Dubins and Freedman (1981) and Roth (1982) have shown that for any man  $m$  with preferences  $u_m$  and beliefs  $\tilde{P}_{-m}$ , it is optimal for  $m$  to submit his true preference list  $P_m$  (which corresponds to  $u_m$ ).

**Theorem.** (*Dubins and Freedman; Roth*) *In the Preference List Submission Problem,*

$$P_m \in \arg \max_{\hat{P}_m \in \mathcal{P}_m} \mathbb{E}[u_m(\mu^M[\hat{P}_m, \tilde{P}_{-m}](m))].$$

This is not the case for women, as they may misrepresent their preferences and get preferable outcomes in some settings (Roth, 1982). A natural way to misrepresent one's preferences is by submitting a *truncated* preferences list.

A truncated preference list is identical to the original one, except that some acceptable partners are declared unacceptable. Denote by  $P_w^k$  the preference list which includes in order only  $w$ 's  $k$  most preferred men, and call this the  $k$ -truncation of her true preference list  $P_w$ . If fewer than  $k$  men are acceptable to  $w$ , then  $P_w^k \equiv P_w$ . Truncation generates a simple tradeoff which is described by the following proposition:

**Proposition.** *Let  $P$  be the preference list profile of all agents in  $\mathcal{M} \cup \mathcal{W}$ . Then  $\mu^M[P_w^k, P_{-w}](w)$  is  $w$ 's least preferred achievable mate under  $P$  with rank  $\leq k$ . Should no such mate exist,  $\mu^M[P_w^k, P_{-w}](w) = w$ .*

The proposition implies that when others' submitted preferences lists are known with certainty it is easy to find a truncation strategy that would match the woman with her most preferred achievable partner, but also that when there exists uncertainty about others' submitted lists truncation may yield each of the three possible results relative to truthful reporting:

1. No effect - when woman  $w$  has truncated below her least preferred achievable mate
2. Improvement - when woman  $w$  truncates above her least preferred achievable mate, and is matched with her least preferred achievable mate above the point of truncation
3. Turning unmatched - when woman  $w$  has an achievable mate, but has over-truncated by truncating above her most preferred achievable mate

Since the realized outcome depends on the realized profile that others submit, each truncation yields a lottery given the beliefs  $\tilde{P}_{-w}$ , and the problem of



choosing the optimal truncation corresponds to choosing the most preferable lottery.

### Optimality of Truncation

Truncation is not the only possible misrepresentation of preferences. A woman could reverse two men in her preference list, list men as acceptable who are in fact unacceptable, drop men from the middle of her list, or use some combination of these. However, under some conditions, truncation is optimal.

The next proposition states that under certainty, women can do no better than to truncate (Roth and Vande Vate, 1991).

**Proposition.** *(Roth and Vande Vate) Suppose woman  $w$  has preferences  $u_w$  and knows others will report preference lists  $P_{-w}$  to MP-DA. Then truncating such that  $\mu^W(w)$  is the last acceptable partner on her list is an optimal strategy for  $w$ .*

Perhaps surprisingly, when a woman has very little information about the preference lists others might report, she again can do no better than to truncate. In order to gain from non-truncation misrepresentations, such as swapping the positions of two men in her reported preference list, a woman must have very specific information about the preference lists others report. Without such information, it is best to leave the men in their correct order. Roth and Rothblum (1999) demonstrate this principle using the following framework.<sup>2</sup>

---

<sup>2</sup>Ehlers (2004) provides weaker conditions, in the same spirit, under which truncation is still optimal.

Let woman  $w$ 's beliefs about reported preference lists of others be represented by  $\tilde{P}_{-w}$ , a random variable taking on values in  $\mathcal{P}_{-w}$ . If  $P_{-w}$  is a preference list profile for agents  $-w$ , define  $P_{-w}^{m \leftrightarrow m'}$  to be the preference list profile in which  $m$  and  $m'$  swap preference lists, and all women swap the positions of  $m$  and  $m'$  in their lists. We say that woman  $w$ 's beliefs are  $(m, m')$ -symmetric if  $\Pr(\tilde{P}_{-w} = P_{-w}) = \Pr(\tilde{P}_{-w} = P_{-w}^{m \leftrightarrow m'})$  for all  $P_{-w} \in \mathcal{P}_{-w}$ . For a subset  $\mathcal{M}' \subseteq \mathcal{M}$ , beliefs  $\tilde{P}_{-w}$  are  $\mathcal{M}'$ -symmetric if they are  $(m, m')$ -symmetric for all  $m, m' \in \mathcal{M}'$ .

**Theorem.** (*Roth and Rothblum*) Suppose  $w$ 's beliefs about reported preference lists of others are  $\mathcal{M}$ -symmetric. Then any preference list  $\hat{P}_w$  she might submit to MP-DA is weakly  $P_w$ -stochastically dominated by some truncation of her true preference list.<sup>3</sup>

Hence, when  $w$  is certain about reported preference lists of her opponents, or when she has extreme, symmetric uncertainty, truncation is optimal.

### The Truncation Problem

Even when truncation is not optimal, we may sometimes wish to restrict the choice set for women to truncations of her true preference list. We define the *Truncation Problem* for woman  $w$  with preferences  $u_w$  and beliefs  $\tilde{P}_{-w}$  on others' submitted preference lists as

$$\max_{k \in \{0, \dots, N\}} \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))].$$

---

<sup>3</sup>  $\hat{P}_w$  is  $P_w$ -stochastically dominated by  $\hat{P}'_w$  iff for any vNM utility function that corresponds to  $P_w$ , the expected utility from submitting  $\hat{P}'_w$  is at least as great as the expected utility from submitting  $\hat{P}_w$ .

### 4.3 Optimal Truncation in Unbalanced Markets

Following Coles and Shorrer (2014) and Ashlagi *et al.* (2013) we consider a setting where each agent draws independently uniformly at random a complete preference list (so that all mates are acceptable).<sup>4</sup> We assume further that for each agent  $i$ ,  $u_i(\cdot)$  is linear in the rank of  $i$ 's match, where being unmatched is treated as rank one below the lowest ranked mate. For a balanced uniform market with  $N$  men and  $N$  women, define

$$k^*(N) \equiv \max_{k \in \{0, \dots, N\}} \left( \arg \max \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))] \right).$$

$k^*(N)$  describes woman  $w$ 's optimal point of truncation, given that the other agents submit their true preference lists. If there are multiple optima, we conservatively select that which involves the least truncation. Coles and Shorrer (2014) prove the following theorem.

**Theorem 4.1.** *Let woman  $w$  have uniform beliefs and preferences linear in rank (or any strictly increasing, convex transformation of such preferences). Then  $\lim_{N \rightarrow \infty} \frac{k^*(N)}{N} = 0$ .*

Theorem 4.1 states that for balanced markets, as the market size grows large, the fraction of the list that an individual woman optimally truncates goes to 100%. The intuition behind this theorem can be gleaned from statistical facts about the most and least preferred achievable mates for women.

In large balanced markets where preferences are uniform, the expected rank

---

<sup>4</sup>While this assumption is not very realistic for real markets, it may serve as an approximation for the behavior of the top tiers in a tiered market. For example, it may be the case that everyone agrees about the composition of the top tier of schools and students, but personal tastes causes the orderings to vary (Ashlagi *et al.*, 2013).

of the most preferred achievable mate of a woman (which is the same as the expected rank of her mate under *WP-DA*) is very low relative to the length of her list; it asymptotes to  $\log N$  (Pittel, 1989). This suggests that a woman may safely truncate a large fraction of her list with little risk of becoming unmatched. Furthermore, the expected rank of a woman's match under *MP-DA* is significantly worse, asymptoting to  $\frac{N}{\log N}$  (Pittel, 1989). In fact, for large markets, Pittel (1992) proved that the worst-off wife will be matched with a husband at the bottom of her list with probability approaching 1. This large gap in a woman's expected most and least preferred achievable mates suggests that not only is it safe to truncate a large fraction of one's list in large markets, but that a woman will also generate gains from such a truncation.

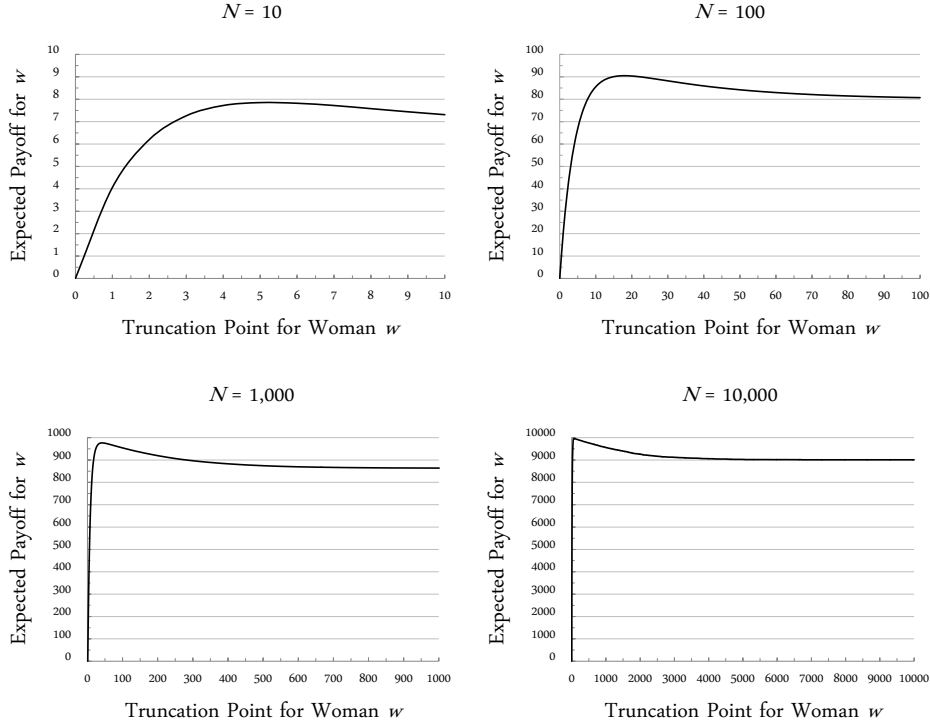
Figure 4.1 presents simulation results for balanced markets of size 10, 100, 1,000 and 10,000. It is clear from the figures that, when all other agents are truthful, the best response of a strategic woman is to submit a (very) short list. It also appears that the gains from truncation may be significant. In a market of size 10,000, the partner rank could potentially be reduced by about 1,000 in expectation (10% of the market size).

The recent paper by Ashlagi *et al.* (2013) implies that the large gap between the best and worst stable partners is a knife-edge case. When markets are even slightly unbalanced, under any stable matching the rank of the mates that an agent on the over demanded side of the market gets is approximately  $\log N$  in expectation, while the other side can expect approximately  $\frac{N}{\log N}$ .<sup>5</sup>

---

<sup>5</sup>The approximations calculated by Ashlagi *et al.* (2013) involve multiplicative constants and describe expected payoffs conditional on an agent being assigned a partner.

Intuitively, these results imply that unbalanced matching markets typically have a “small core.” In turn this suggests that submitting a long list may constitute an optimal strategy in unbalanced uniform markets. In light of their findings, we inspect strategic behavior in the unbalanced setting.



**Figure 4.1: *Simulation Results for Truncation Payoffs.*** The graphs display  $(N + 1)$ — an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, N\}$  and submitting these preferences to MP-DA. Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. Markets are of size 10, 100, 1,000 and 10,000.

#### 4.3.1 The Case of More Women Than Men

Given  $L \geq 0$ , for a market with  $N$  men and  $N + L$  women define

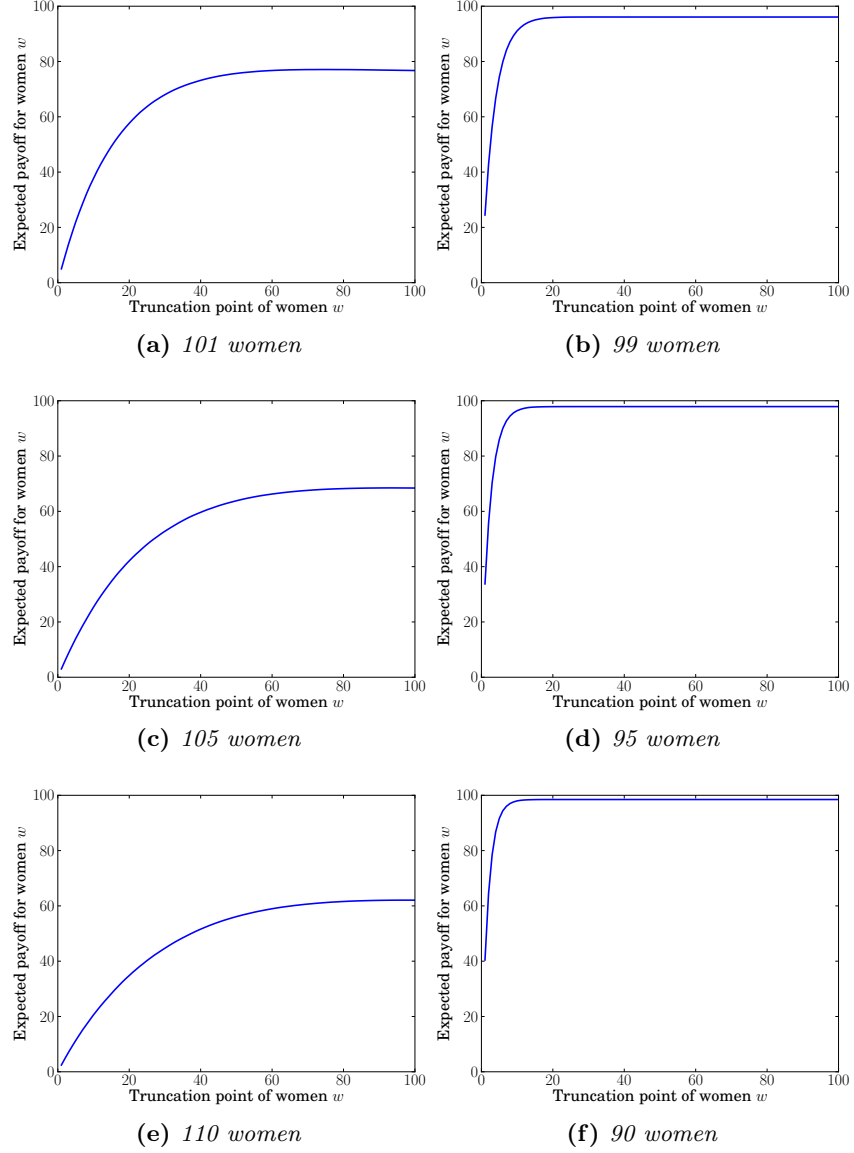
$$k_*(N, L) \equiv \min \left\{ \left( \arg \max_{k \in \{0, \dots, N\}} \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))] \right) \right\}.$$

$k_*(N, L)$  describes woman  $w$ 's optimal point of truncation, given that the other agents submit their true preference lists. If there are multiple optima, we conservatively select that which involves the most truncation. Note that for  $L = 0$ ,  $k_*(N, 0) \leq k^*(N)$  so the results of Theorem 4.1 apply to  $k_*(N, 0)$ . We now have the following theorem, which constitutes a partial converse of Theorem 4.1. The theorem shows that the intuition from Ashlagi *et al.* (2013) extends to the incomplete information setting when women are over demanded.

**Theorem 4.2.** *Given  $L > 0$ , consider a market with  $N$  men and  $N + L$  women. Let woman  $w$  have uniform beliefs and preferences linear in rank (or any strictly increasing, concave transformation of such preferences). Then  $\frac{k_*(N, L)}{N} \geq \frac{L}{L+1}$  so  $\lim_{N \rightarrow \infty} \frac{k_*(N, L)}{N} \geq \frac{L}{L+1}$ . In particular  $\frac{k_*(N, L)}{N} \geq \frac{1}{2}$  and  $\lim_{N \rightarrow \infty} \frac{k_*(N, L)}{N} \geq \frac{1}{2}$ .*

*Proof.* Recall that a truncation by woman  $w$  could lead to one of three results:

1. No effect: woman  $w$  has truncated below her least preferred achievable mate
2. Improvement: woman  $w$  truncates above her least preferred mate, and is matched with her least preferred achievable mate above the point of truncation
3. Becoming unmatched: woman  $w$  has an achievable mate, but has over-truncated, truncating above her most preferred achievable mate



**Figure 4.2: Simulation Results for Truncation Payoffs.** The graphs display 101– an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, 100\}$  and submitting these preferences to MP-DA. Preference lists of the other agents are uniformly random, selected from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws. All markets have 100 men, and the number of women varies between 90, 95, 99, 101, 105 and 110.

In a balanced market, truncation worsened a woman's outcome only when *MP-DA* terminated with one man exhausting his list. But when women outnumber men, truncation may cause *MP-DA* to terminate when a previously unmatched woman receives an offer. Note that when agents are truthful, at least one woman will have received no proposals prior to the truncation. Of course improvement upon truncation is possible only if  $w$  does not end up single.

Using the *principle of deferred decisions*, it is easy to see that conditional on a truncation making a difference, the probability of improvement is less than  $\frac{1}{L+1} \leq \frac{1}{2}$ . To show this, recall that unmatched women have not received any proposals. Hence, it follows from the principle of deferred decisions and symmetry that, following truncation, any future proposal is at least as likely to be directed at these women as to  $w$ . The algorithm terminates only when such a proposal happens.

Now consider the marginal benefit to  $w$  from omitting the lowest-ranked man from a list of length  $m + 1$ . The most  $w$  can hope for is an improvement in her match of  $m$  ranks (from  $m + 1$  to the top).<sup>6</sup> If this omission instead leaves her unmatched, she drops  $N - m$  ranks (from  $m + 1$  to  $N + 1$ ). Since the probability of becoming unmatched conditional on truncation having any effect is at least  $\frac{L}{1+L} \geq \frac{1}{2}$ , the expected gain cannot be positive if  $m < \frac{L}{L+1}N$ . Hence, the optimal list length for  $w$  is at least  $\frac{L}{L+1}N \geq \frac{1}{2}N$ .  $\square$

The left panels of Figure 4.2 illustrate our findings. We simulated markets with 100 men and 99, 95 and 90 women. In each market we generate

---

<sup>6</sup>In fact, the gain is  $\frac{m+1}{2}$  in expectation - the expected rank of the remaining partners.



independently and at random a full preference list for each agent. We then calculate an individual woman’s payoff, given that all other agents submit their true lists, for each possible level of truncation. Payoffs are depicted by  $101 - \text{partner rank}$ , and we report the average result over 100,000 iterations.

The simulations support the findings of Theorem 4.2, as the optimal list length in all three markets is greater than 50. Indeed, the optimal list lengths are higher than 50, 84 and 91, the respective lower bounds the theorem indicates. In contrast to Figure 4.1, the balanced market case, it is almost impossible to detect the peak of the graphs. That is, not only should women submit long lists, but there is little to gain by truncating optimally. Note that as one would expect, women do worse as the competition on their side increases.

A few points related to Theorem 4.2 deserve attention. First, under uniform beliefs, Roth and Rothblum’s optimality theorem applies whether the market is balanced or not. This implies that the truncation strategies described in Theorems 4.1 and 4.2 are the best overall strategies, not just the optimal truncation strategies. Hence, we have a natural metric for the “distance” between the optimal strategy and truthfulness. The importance of Theorem 4.1 is in showing that the best response to straightforward behavior of others could be “far” from truthful, and so provides an important qualification to the literature which finds truthful reporting to be close to optimal (Roth and Peranson, 1999; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Lee, 2011; Ashlagi *et al.*, 2013). Theorem 4.2 qualifies this previous finding. We show that our example relies heavily on the fact that the number of women is not larger than the number of men.

A second point worth noting is that our choice of  $k_*(N, L)$  was a conservative one. We could have instead chosen to state the theorem using  $k^*(N, L) \equiv \max \left\{ \left( \arg \max_{k \in \{0, \dots, N\}} \mathbb{E}[u_w(\mu^M[P_w^k, \tilde{P}_{-w}](w))] \right) \right\}$ , breaking ties in favor of less truncation as in the definition of  $k^*(N)$ , and the theorem would of course still hold (since  $k^*(N, L) \geq k_*(N, L)$  by definition).

Finally, in contrast to most of the results in this strand of the literature, our theorem is *not* a “large market result”; our result holds for any  $N$  and  $L$ . That is, manipulation opportunities are minimal (in the sense of distance from truthful submission) whenever women outnumber men, for unbalanced markets of any size. However, a simple corollary of our result is that as imbalance in a market increases, manipulation opportunities vanish altogether.

**Corollary 4.1.** *Given a sequence  $\{L_N\}$  with  $\lim_N L_N = \infty$  and a sequence of uniform markets with  $N$  men and  $N + L_N$  women, if woman  $w$  has uniform beliefs and preferences linear in rank (or any strictly increasing, concave transformation of such preferences), then  $\lim_{N \rightarrow \infty} k_*(N, L_N) = 1$ .*

The simulation results presented in Figure 4.2 are consistent with the results of Ashlagi *et al.* (2013). In contrast to the relatively large gain that a woman may be able to realize by truncating in a balanced market, when there are more women than men the graph of the expected payoff is much flatter between the optimum and truthful reporting. It is also true that in this case truncation is relatively risky. One can see from Figure 4.2 that, for example, in a market with 101 women and 100 men, submitting a list of length 30 exposes a woman to a significant risk of remaining unmatched. The

following theorem formalizes this observation by providing a lower bound on the probability of becoming unmatched following a relatively conservative truncation.

**Theorem 4.3.** *Fix  $L > 0$  and  $\delta \in (0, 1)$ . For  $N$  large enough, in a uniform market with  $N$  men and  $N + L$  women, if all other agents report truthfully and woman  $w$  submits a truncation list of length less than  $\delta N$ , she will be unmatched with probability at least  $\frac{.49+L}{N+L}$ .*

*Proof.* From Pittel (1992, Theorem 6.2) we know that in a balanced uniform market of size  $N$ , the probability that the worst-off woman gets a mate ranked worse than  $\delta N$  approaches 1. This probability only increases when there are more women than men (Kelso and Crawford, 1982). Now consider a woman truncating her list shorter than  $\delta N$  while all other agents are truthful. From Pittel's theorem, for large  $N$  the probability that this truncation makes a difference is at least  $\frac{1}{N+L+1}$ . For large enough  $N$ , this expression is greater than  $\frac{.999}{N}$ . Conditional on the truncation making a difference, using the principle of deferred decisions, the resulting chain of rejections is at least as likely to terminate with a proposal to a woman that did not receive any proposals until  $w$  divorced her partner, as it is to return to  $w$ . This implies that even in the event that  $w$  is matched when she (and all others) report truthfully, by truncating her list to a size smaller than  $\delta N$  she raises her probability of being unmatched by at least  $\frac{L}{1+L} \times \frac{.999}{N} > \frac{.49}{N}$ . Multiplying both sides of the inequality by the probability of  $w$  being matched if she reports truthfully, and adding the probability of  $w$  being unmatched if she reports

truthfully, produces the lower bound:

$$\frac{L}{N+L} + \left(1 - \frac{L}{N+L}\right) \frac{.49}{N} = \frac{L + .49}{N+L}.$$

□

Symmetry implies that  $w$  will remain unmatched with probability  $\frac{L}{N+L}$  no matter what (full) list she reports. We are interested in the increase in probability of being left unmatched relative to truthful reporting  $\left(\frac{.49}{N+L}\right)$ . While the increase does not appear to be large at first glance, several facts must be taken into account. First, this is a lower bound, and there is no good reason to suspect that it is tight. Moreover, the lower bound for the increase in probability is of the same order of magnitude as the probability of remaining unmatched under truthful reporting. Second, the degree of truncation of  $w$  may be minimal. The theorem allows  $w$  to submit 99% of her list and the results will still hold. A third point is that these results should be compared with the opposite case, where men outnumber women. This is exactly what we do in the next section.

### 4.3.2 The Case of Fewer Women Than Men

The results of Ashlagi *et al.* (2013) regarding the small core apply regardless of the direction of the imbalance. That is, no matter the size or direction of the imbalance, the expected potential improvement is small. One might therefore suspect that when men outnumber women, an analog to Theorem 4.2 would apply and room for manipulation would again be small. This, however, does not appear to be the case. The simulation results presented in

the right panel of Figure 4.2 indicate that when there are fewer women than men, the optimal level of truncation may still be significant.

The figure depicts (simulated) truncation payoffs for men in markets with 100 men and 101, 105 and 110 women. In contrast to the case where women outnumber men, in this case the peaks of all three graphs involve lists of length shorter than 31.

Comparing the right and left panels of Figure 4.2, three additional facts stand out. The first is that women do much better when the balance tips slightly in their favor: payoffs with 99 women and 100 men are much higher than when there are 101 women. This difference becomes starker as the imbalance increase. This corroborates the findings of Ashlagi *et al.* (2013) in the case when  $w$  reports truthfully. Also salient is that even though optimal truncation may be far from the truth, such a manipulation increases payoffs only minimally. This too could be deduced from their paper. The third salient feature is not a direct consequence of their findings (though it is related to techniques used in their proofs). The simulations suggest that truncation is “safer” for women when they are over-demanded. That is, when there are more men than women, women may submit relatively short lists without facing a large risk of becoming unmatched, even if there is little gain from doing so.

The next theorem shows that this third fact holds more generally.

**Theorem 4.4.** *Fix  $L \geq 0$ . For a uniform market with  $N + L$  men and  $N$  women, a woman that submits a truncation containing more than  $L + (2 + a) \log^2 N$  men will be matched with probability at least  $1 - O(N^{-c(a)})$ , where*

$c(a) = 2a \left[ 3 + (4a + 9)^{\frac{1}{2}} \right]^{-1}$ . In particular a women that submits a truncated list of more than  $L + 10 \log^2 N$  men will be unmatched with probability at most  $O\left(\frac{1}{N^2}\right)$ .

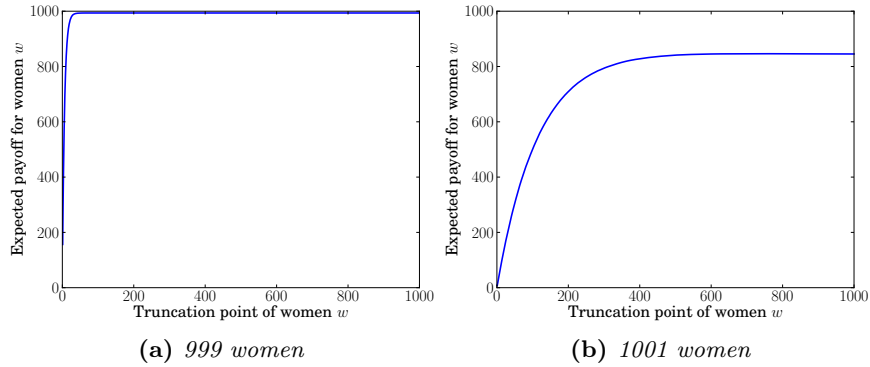
*Proof.* For notational simplicity we provide the proof for the case of  $L + 10 \log^2 N$ , but analogous arguments would apply for all other cases. The proof has two steps. First, recall that in a market that uses *MP-DA*, adding men to the market makes the other men weakly worse-off and women weakly better-off (Kelso and Crawford, 1982). Second, from Pittel (1992, Theorem 6.1) we know that in a balanced market of size  $N$ , submitting a truncated list with  $10 \log^2 N$  men ensures being matched with probability  $1 - O\left(\frac{1}{N^2}\right)$ . In a market with all the women and an arbitrary subset of the men containing  $N$  agents, by submitting a truncated list with  $L + 10 \log^2 N$  men, a woman submits a list containing, at least, her most preferred  $10 \log^2 N$  men in the subset in order. By Pittel's theorem, this ensures that the woman is matched with probability at least  $1 - O\left(\frac{1}{N^2}\right)$ . But the first point ensures that by adding the other  $L$  men to the market all women are weakly better-off. In particular, no woman that would have been matched in the smaller market can become unmatched.  $\square$

*Remark.* The statement of Theorem 4.4 is intentionally silent on the strategies of women  $-w$ . The proof shows that the statement holds when all other women are truthful. But the proof also holds whenever other women use truncation strategies, or any anonymous strategies. The logic is simple: truncation by other women only increases  $w$ 's probability of being matched given any list she submits.

Intuition for Theorem 4.4 may come from considering markets with large imbalances. Consider, for example, uniform markets with  $N$  women and  $(1 + \lambda)N$  men, for positive  $\lambda$ . In these markets, *MP-DA* terminates only after  $\lambda N$  men have proposed to all of the women. Since preferences are independent, this implies (using the “principle of deferred decisions”) that even in the men optimal stable matching, a woman is matched with a man high on her list with high probability (her expected partner rank is lower than  $\frac{N}{\lambda N} = \frac{1}{\lambda}$ ).

Rather than being an analog, Theorem 4.4 stands in sharp contrast to Theorem 4.3. The ratio between the length of the lists described in Theorem 4.4 and the ones described in Theorem 4.3 approaches 0 as  $N$  grows large (since the lists from Theorem 4.4 are much shorter). Yet the ratio of the increases in the probability of becoming unmatched approaches infinity if the (short) list in the setting of Theorem 4.4 is chosen to be sufficiently long (e.g.  $11 \log^2 N$ ).

To illustrate Theorem 4.4, we present additional simulation evidence. We simulated a market with 1000 men and 999 women, and estimated the returns to truncation for a woman  $w$  given that all other agents are truthful, reporting average results over 100,000 iterations. The results are summarized in the left panel of Figure 4.3. While difficult to observe with the naked eye, the maximum is attained at 89, so that in terms of list length, the best response is still “far” from truthful reporting.



**Figure 4.3: Simulation Results for Truncation Payoffs.** In markets with 1000 men and 999 or 1001 women, the graph displays 1001— an individual woman’s expected partner rank from truncating her list at each point  $k \in \{0, \dots, 1000\}$  and submitting these preferences to MP-DA. Preference lists of the other agents are selected, uniformly at random, from the set of all possible full length preference list profiles, and payoffs are averaged over 100,000 draws.

## 4.4 Other Aspects Impacting the Optimal Level of Truncation

Coles and Shorrer (2014) provide several comparative statics for the optimal level of truncation in the case of balanced markets. We demonstrate that these hold in the unbalanced case as well.

### 4.4.1 Truncation and Risk Aversion

As discussed previously, truncation is a risky strategy. Compared to truthful reporting, truncation may offer some benefit, but over-truncating can lead to large losses depending on the profile of preferences that is submitted by others. One might expect agents with more conservative attitudes toward risk to shy away from this proposition. In this section, we formalize this



intuition.

We consider a general setting, with arbitrary preferences for woman  $w$  and beliefs about reported preferences of others. Let  $\psi(\cdot)$  be any strictly increasing, concave transformation. We claim that for any beliefs about others, woman  $w$  with preferences  $u_w(\cdot)$  will truncate more than a woman  $w_\psi$  who has identical beliefs, but preferences given by  $\psi(u_w(\cdot))$ .

We fix  $w$ 's preferences to be  $u_w(\cdot)$ , and define the shorthand

$$v(k, P_{-w}) \equiv u_w(\mu^M[P_w^k, P_{-w}](w)),$$

$w$ 's payoff from submitting truncated preference list  $P_w^k$ . Now define

$$v_\psi(k, P_{-w}) \equiv \psi(u_w(\mu^M[P_w^k, P_{-w}](w))),$$

the payoff from submitting truncated preference list  $P_w^k$  for a woman  $w_\psi$  with preferences  $\psi(u_w(\cdot))$ .

The following theorem states that if  $w$  prefers truncating less to more, then  $w_\psi$  definitely prefers truncating less to more.

**Theorem 4.5.** *Let  $\tilde{P}_{-w}$  be any random variable distributed over  $\mathcal{P}_{-w}$ . Then  $\forall k \in \{1, \dots, N-1\}, \forall t \in \{1, \dots, N-k\}$  we have*

$$\begin{aligned} \mathbb{E}[v(k, \tilde{P}_{-w})] &\leq \mathbb{E}[v(k+t, \tilde{P}_{-w})] &\Rightarrow \\ \mathbb{E}[v_\psi(k, \tilde{P}_{-w})] &\leq \mathbb{E}[v_\psi(k+t, \tilde{P}_{-w})]. \end{aligned}$$

*Furthermore, if i)  $\psi(\cdot)$  is strictly concave, and ii) under  $\tilde{P}_{-w}$ , each man is achievable for  $w$  with positive probability, then the second inequality is strict.*

We can now use Theorem 4.5 to sort optimal truncation points based on

degree of concavity.

**Corollary 4.2.** *Let  $k_i^l$  be the minimum optimal truncation point (by rank) and let  $k_i^h$  be the maximum optimal truncation point for woman  $i \in \{w, w_\psi\}$ . Then  $k_w^l \leq k_{w_\psi}^l$  and  $k_w^h \leq k_{w_\psi}^h$ . Furthermore, if conditions i) and ii) from Theorem 4.5 hold, then  $k_w^h \leq k_{w_\psi}^l$ .*

We omit the proofs, as they are straightforward analogs of the proofs of Theorem 5 and Corollary 1 in Coles and Shorrer (2014). The key insight in the analysis is the interpretation of truncation as a risky lottery, and then mapping the additional risk associated with incremental truncation to an *extra lottery* a woman must face. If a woman doesn't like to face the extra lottery, then certainly a woman with more concave preferences will not want to face it. Note that despite pertaining to risk aversion, the results in this section do not restrict the structure of  $u_w(\cdot)$  in any way. For example, we do not require  $u_w(\cdot)$  to be “concave.” Rather, it is the *relative* concavity that is crucial.

This result can offer advice to a market designer. If she wishes to see long lists, for example since her objective is to maximize the number of matches, a market designer may wish to choose the less risk averse side to be the “proposers” in the Deferred Acceptance Algorithm. If the two sides of the market are identical in all regards except for their risk preferences, the more risk averse side will be less likely to truncate, even if manipulations increase their expected partner rank. Lower levels of truncation will increase the number of realized matches, and consequently, reduce the number of participants left unmatched. However, in making this choice, the market

designer should take other market features into consideration as well, as we demonstrate in the next section.

#### 4.4.2 Truncation and Correlated Preferences

Coles and Shorrer (2014) provided theoretical and empirical evidence that, in the balanced setting, correlation in preferences of agents on one side of the market reduces their incentive to truncate. In this section, we show that their findings generalize to unbalanced markets.

We consider first the case of perfectly correlated preferences on women's side of the market. In this case, there exist a unique stable matching, and so women have no incentive to truncate their lists at all when all others report truthfully. If women are uncertain about men's preferences, truncation may only lead them to a worse outcome, provided that others are truthful.

While perfect correlation and independence are easy to model, partial correlation may appear in many forms. In this paper, we focus on one simple such form. Consider the Preference List Submission Problem for woman  $w$  with preferences  $u_w$  and beliefs  $\tilde{P}_{-w}$  about reported preference lists of opponents. Let  $p(\cdot, \cdot)$  be the probability mass function for  $w$ 's beliefs. That is,

$$p(P_{\mathcal{M}}, P_{\mathcal{W} \setminus \{w\}})$$

gives the likelihood that the men will report preference lists  $P_{\mathcal{M}}$  and women  $\mathcal{W} \setminus \{w\}$  will report preference lists  $P_{\mathcal{W} \setminus \{w\}}$ . Define the marginal probability over men's preference profiles by  $p^M(\cdot)$ .

Given  $p(\cdot, \cdot)$ , define beliefs  $p^C(\cdot, \cdot)$  by

$$p^C(P_{\mathcal{M}}, P_{\mathcal{W} \setminus \{w\}}) \equiv \begin{cases} p^M(P_{\mathcal{M}}) & \text{if } P_{\hat{w}} = P_w \quad \forall \hat{w} \in \mathcal{W} \setminus \{w\} \\ 0 & \text{otherwise} \end{cases}.$$

$p^C(\cdot, \cdot)$  is the distribution that preserves the marginal distribution over men's preferences  $p^M(\cdot)$ , but where the other women share the preferences of  $w$ .

Define beliefs  $p^\alpha(\cdot)$  by

$$p^\alpha(P_{-w}) \equiv (1 - \alpha)p(P_{-w}) + \alpha p^C(P_{-w}).$$

Hence, as  $\alpha$  varies from 0 to 1,  $p^\alpha$  ranges from  $p$  to  $p^C$ . The marginal distribution over men's preferences remains fixed, while the correlation of women's preferences steadily increases (the distribution remains constant if  $p = p^C$ ).

The set of optimal truncation points for woman  $w$  with preferences  $u_w$  and beliefs indexed by  $\alpha$  is given by

$$k^*(\alpha, p, u_w) \equiv \arg \max_{k \in \{0, \dots, N\}} \mathbb{E}_{p^\alpha}[v(k, \tilde{P}_{-w})].$$

Notice that since the choice set is finite,  $k^*(\cdot, \cdot)$  will be non-empty.

Let  $k^h(\alpha, p, u_w) = \max[k^*(\alpha, p, u_w)]$  and  $k^l(\alpha, p, u_w) = \min[k^*(\alpha, p, u_w)]$ , the optimal choices involving the least and most truncation respectively.

The following proposition states that for any preferences  $u_w$  and beliefs  $p$ , as we increase the degree of correlation  $\alpha$ , woman  $w$  should truncate less.

**Proposition 4.1.** *Let  $\alpha, \alpha' \in [0, 1]$  with  $\alpha' > \alpha$ . Then  $k^l(\alpha', p, u_w) \geq k^l(\alpha, p, u_w)$  and  $k^h(\alpha', p, u_w) \geq k^h(\alpha, p, u_w)$ .*

The proof of the proposition is analogous to the proof of Proposition 2 in Coles and Shorrer (2014), and is therefore omitted.

The anticipated level of correlation in the environment might influence the advice a market designer can offer participants. If correlation is high, the designer can safely advise participants to report truthfully, and it is in their best interest to do so. With low correlation (sufficiently heterogeneous preferences), players may anticipate gains from truncation, which if acted on, could lead to unstable matching.

## 4.5 Conclusion

In this paper, we study optimal strategic behavior in unbalanced one-to-one matching markets, where matchings are determined by the Deferred Acceptance Algorithm and agents have incomplete information about the preferences of others. We focus on truncation strategies, which are attractive for agents as they are simple and always weakly increase the probability of being matched with more-preferred mates. From a computational perspective, this reduces significantly the dimensions of the strategy space, allowing us to use simulations to pinpoint optimal behavior. This restriction also induces a natural metric on the extent of manipulation: the shorter the lists submitted, the further they are from truthfulness. This allows us to make relative statements about optimal list.

The main innovation of this paper is in studying the effect of imbalance in the number of agents on the two sides of a market on their potential for manipulation. We study a stylized setting which we term a uniform

market, and find that the degree of manipulation observed in this setting critically depends on the direction of the imbalance. When women are on the long side of the market (there are more women than men), we find that the incentives for women to manipulate are significantly diminished compared to a balanced market. This finding is consistent with the intuition of Ashlagi *et al.* (2013), who find that the expected gap between an agent’s highest and lowest achievable mates is small in unbalanced uniform markets.

By contrast, when men outnumber women, we provide evidence that a woman’s best response to truthful behavior by others involves a significant degree of truncation. This finding qualifies results that suggest opportunities for manipulation in such settings are minimal (e. g. in terms of potential gain in utility (Lee, 2011)). We further show that truncation is safe when women are on the short side (more men than women) but not when they are on the long side.

We also provide comparative statics regarding the extent of manipulation, regardless of the direction of size of a market imbalance. When women are more risk averse, they should be less aggressive in their degree of truncation. Correlation in womens’ preferences also reduces their incentive to truncate.

Matching mechanisms based on the Deferred Acceptance Algorithm are used extensively in a variety of entry level labor markets (Roth, 1990) and in school choice (Pathak and Sonmez, 2008). One advantage of DAA is that it induces truthful reporting as a dominant strategy for one side of the market (Roth, 1982; Dubins and Freedman, 1981). This alone is an argument market designers have used to decide which side will be the “proposing” one (Roth, 1990).

In addition to shedding light on strategic behavior in unbalanced markets generally, our paper introduces a new factor that might be considered when selecting the proposing side: direction of imbalance. By selecting the over-demanded side to propose, potential for strategic manipulation is minimized. Selecting the over-demanded side to receive offers leaves room for significant, safe manipulation. While simplistic and stylized in many respects, our result is a first effort to extend the logic market designers rely on in choosing the proposing side. Future work should find more general environments in which the extent of manipulation may be compared, and explore the interaction between the different forces that determine the incentives to manipulate.

# References

- ABDULKADIROĞLU, A., PATHAK, P. A. and ROTH, A. E. (2005). The new york city high school match. *American Economic Review*, pp. 364–367.
- AGARWAL, V. and NAIK, N. Y. (2000). Multi-period performance persistence analysis of hedge funds. *Journal of Financial and Quantitative Analysis*, **35** (03), 327–342.
- ANTRÀS, P., CHOR, D., FALLY, T. and HILLBERRY, R. (2012). *Measuring the upstreamness of production and trade flows*. Tech. rep., National Bureau of Economic Research.
- ARROW, K. J. (1965). *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Jahnssonin Säätiö.
- (1971). *Essays in the Theory of Risk Bearing*. Chicago: Markham.
- (1972). The value of and demand for information. *Decision and organisation*. Londres: North-Holland.
- and LEVHARI, D. (1969). Uniqueness of the internal rate of return with variable life of investment. *The Economic Journal*, **79** (315), pp. 560–566.
- ARTZNER, P. (1999). Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal*, **3** (2), 11–25.
- , DELBAEN, F., EBER, J.-M. and HEATH, D. (1999). Coherent measures of risk. *Mathematical finance*, **9** (3), 203–228.
- ASHLAGI, I., KANORIA, Y. and LESHNO, J. D. (2013). Unbalanced random matching markets. In *ACM Conference on Electronic Commerce*, pp. 27–28.
- and KLIJN, F. (2012). Manipulability in matching markets: Conflict and coincidence of interests. *Social Choice and Welfare*, **39**, 23–33.



- ATHEY, S. and LEVIN, J. (2001). The value of information in monotone decision problems.
- ATKINSON, A. B. (1970). On the measurement of inequality. *Journal of economic theory*, **2** (3), 244–263.
- (1983). *The economics of inequality*. Clarendon Press Oxford.
- (1987). On the measurement of poverty. *Econometrica*, **55** (4), pp. 749–764.
- AUMANN, R. J. and SERRANO, R. (2008). An economic index of riskiness. *Journal of Political Economy*, **116** (5), pp. 810–836.
- AZEVEDO, E. M. and LESHNO, J. D. (2012). A supply and demand framework for two-sided matching markets. In *MATCH-UP 2012: the Second International Workshop on Matching Under Preferences*, p. 124.
- BANZHAF III, J. F. (1964). Weighted voting doesn’t work: A mathematical analysis. *Rutgers L. Rev.*, **19**, 317.
- BARRO, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, **121** (3), 823–866.
- (2007). *Rare disasters, asset prices, and welfare costs*. Tech. rep., National Bureau of Economic Research.
- BENOÎT, J.-P. and OK, E. A. (2007). Delay aversion. *Theoretical Economics*, **2** (1), 71–113.
- BINSWANGER, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural india. *The Economic Journal*, pp. 867–890.
- BLACKWELL, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, **24** (2), pp. 265–272.
- BØHREN, Ø. and HANSEN, T. (1980). Capital budgeting with unspecified discount rates. *The Scandinavian Journal of Economics*, **82** (1), pp. 45–58.
- BONACICH, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, pp. 1170–1182.
- CABRALES, A., GOSSNER, O. and SERRANO, R. (2012). The appeal of information transactions. *Working paper*.

- , — and — (2013). Entropy and the value of information for investors. *The American Economic Review*, **103** (1), 360–377.
- , — and — (2014). The inverse demand for information and the appeal of information transactions. *Working paper*.
- CALDARELLI, G. and CAPOCCI, A. (2000). Beauty and distance in the stable marriage problem. *Physica A: Statistical Mechanics and its Applications*, **300**, 325–331.
- COHEN, A. J. and HARCOURT, G. C. (2003). Retrospectives: Whatever happened to the cambridge capital theory controversies? *Journal of Economic Perspectives*, pp. 199–214.
- COLES, P., CAWLEY, J., LEVINE, P., NIEDERLE, M., ROTH, A. E. and SIEGFRIED, J. J. (2010). The job market for new economists: A market design perspective. *Journal of Economic Perspectives*, **24** (4), 187–206.
- , KUSHNIR, A. and NIEDERLE, M. (2013). Preference signaling in matching markets. *American Economic Journal: Microeconomics*, **5** (2), 99–134.
- and SHORRER, R. (2014). Optimal truncation in matching markets. *Games and Economic Behavior*, **87**, 591–615.
- DEBREU, G. (1972). *Theory of value: An axiomatic analysis of economic equilibrium*, vol. 17. Yale University Press.
- DIEWERT, W. E. (1998). Index number issues in the consumer price index. *The Journal of Economic Perspectives*, pp. 47–58.
- DOHMEN, T., FALK, A., HUFFMAN, D., SUNDE, U., SCHUPP, J. and WAGNER, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, **9** (3), 522–550.
- DUBINS, L. and FREEDMAN, D. (1981). Machiavelli and the gale-shapley algorithm. *American Mathematical Monthly*, **88**(7), 485–494.
- DUNCAN, O. D. and DUNCAN, B. (1955). A methodological analysis of segregation indexes. *American sociological review*, pp. 210–217.
- DZIERZAWA, M. and OMÉRO, M. J. (2000). Statistics of stable marriages. *Physica A*, **287**, 321–333.

- ECHENIQUE, F. and FRYER JR., R. G. (2007). A measure of segregation based on social interactions. *The Quarterly Journal of Economics*, pp. 441–485.
- EHLERS, L. (2004). In search of advice for participants in matching markets which use the deferred-acceptance algorithm. *Games and Economic Behavior*, **48** (2), 249–270.
- (2008). Truncation strategies in matching markets. *Mathematics of Operations Research*, **33** (2), 327–335.
- and MASSÓ, J. (2007). Incomplete information and singleton cores in matching markets. *Journal of Economic Theory*, **136**, 587–600.
- EKERN, S. (1981). Time dominance efficiency analysis. *The Journal of Finance*, **36** (5), 1023–1033.
- FAMA, E. F. (1965). The behavior of stock-market prices. *Journal of business*, pp. 34–105.
- FEATHERSTONE, C. and NIEDERLE, M. (2011). School choice mechanisms under incomplete information: an experimental investigation. *Harvard Business School, Unpublished manuscript*.
- FEATHERSTONE, C. R. and MAYEFISKY, E. (2010). Stability and deferred acceptance: Strategic behavior in two-sided matching, working paper.
- FISHER, I. (1930). The theory of interest. *New York: Kelley, Reprint of the 1930 Edition*.
- FLÜCKIGER, Y. and SILBER, J. (1999). *The measurement of segregation in the labor force*. Springer Science & Business Media.
- FOSTER, D. P. and HART, S. (2009). An operational measure of riskiness. *Journal of Political Economy*, **117** (5), pp. 785–814.
- and — (2013). A wealth-requirement axiomatization of riskiness. *Theoretical Economics*, **8** (2), 591–620.
- FOSTER, J. E. and MITRA, T. (2003). Ranking investment projects. *Economic Theory*, **22**, 469–494.
- FRANKEL, D. M. and VOLIJ, O. (2011). Measuring school segregation. *Journal of Economic Theory*, **146** (1), 1–38.

- GABAIX, X. (2008). Variable rare disasters: A tractable theory of ten puzzles in macro-finance. *The American Economic Review*, pp. 64–67.
- GALE, D. and SHAPLEY, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, **69**, 9–15.
- and SOTOMAYOR, M. (1985). Some remarks on the stable matching problem. *Discrete Applied Mathematics*, **11** (3), 223 – 232.
- HADAR, J. and RUSSELL, W. R. (1969). Rules for ordering uncertain prospects. *The American Economic Review*, **59** (1), pp. 25–34.
- HANOCH, G. and LEVY, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, **36** (3), pp. 335–346.
- HART, S. (2011). Comparing risks by acceptance and rejection. *Journal of Political Economy*, **119** (4), pp. 617–638.
- HARVEY, C. R. and SIDDIQUE, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, **55** (3), 1263–1295.
- HERFINDAHL, O. C. (1950). *Concentration in the steel industry*. Ph.D. thesis, Columbia University.
- HODGES, S. (1998). *A generalization of the Sharpe ratio and its applications to valuation bounds and risk measures*. Financial Options Research Centre, Warwick Business School, University of Warwick.
- HOLT, C. A. and LAURY, S. K. (2002). Risk aversion and incentive effects. *American economic review*, **92** (5), 1644–1655.
- HOMM, U. and PIGORSCH, C. (2012a). Beyond the sharpe ratio: An application of the aumann–serrano index to performance measurement. *Journal of Banking & Finance*, **36** (8), 2274–2284.
- and — (2012b). An operational interpretation and existence of the Aumann-Serrano index of riskiness. *Economics Letters*, **114** (3), 265 – 267.
- HOROWITZ, J. K. (1992). Comparative impatience. *Economics Letters*, **38** (1), 25 – 29.
- HUTCHENS, R. (2004). One measure of segregation. *International Economic Review*, **45** (2), 555–578.

- (2015). Symmetric measures of segregation, segregation curves, and blackwell's criterion. *Mathematical Social Sciences*, **73**, 63–68.
- IMMORLICA, N. and MAHDIAN, M. (2005). Marriage, honesty, and stability. *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*.
- JEWITT, I. (2007). Information order in decision and agency problems. *Nuffield College*.
- KADAN, O. and LIU, F. (2014). Performance evaluation with high moments and disaster risk. *Journal of Financial Economics*, **113** (1), 131–155.
- KANE, A. (1982). Skewness preference and portfolio choice. *Journal of Financial and Quantitative Analysis*, **17** (01), 15–25.
- KAT, H. M. and BROOKS, C. (2001). The statistical properties of hedge fund index returns and their implications for investors. *Cass Business School Research Paper*.
- KELSO, J., ALEXANDER S. and CRAWFORD, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica*, **50** (6), pp. 1483–1504.
- KNUTH, D. E. (1976). *Mariages stable et leurs relations avec d'autres problèmes combinatoires: Introduction à l'analyse mathématique des algorithmes*. Les Presses de l'Université de Montréal.
- KOJIMA, F. (2006). Mixed strategies in games of capacity manipulation in hospital-intern markets. *Social Choice and Welfare*, **27**, 25–28.
- and PATHAK, P. (2009). Incentives and stability in large two-sided matching markets. *American Economic Review*, **99**, 608–627.
- KONISHI, H. and ÜNVER, U. (2006). Games of capacity manipulation in hospital-intern markets. *Social Choice and Welfare*, **27** (1), 3–24.
- KRAUS, A. and LITZENBERGER, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, **31** (4), 1085–1100.
- KULLBACK, S. and LEIBLER, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, **22** (1), 79–86.
- LAMBSDORFF, J. G. (2007). The methodology of the corruption perceptions index 2007.

- LASSO DE LA VEGA, C. and VOLIJ, O. (2014). Segregation, informativeness and lorenz dominance. *Social Choice and Welfare*, **43** (3), 547–564.
- LAVALLE, I. H. (1968). On cash equivalents and information evaluation in decisions under uncertainty: Part i: Basic theory. *Journal of the American Statistical Association*, **63** (321), pp. 252–276.
- LEE, R. S. and SCHWARZ, M. (2012). Interviewing in two-sided matching markets, working paper.
- LEE, S. (2011). Incentive compatibility of large centralized matching markets, working paper.
- LEHMANN, E. (1988). Comparing location experiments. *The Annals of Statistics*, pp. 521–533.
- LESHNO, M. and SPECTOR, Y. (1992). An elementary proof of blackwell’s theorem. *Mathematical Social Sciences*, **25** (1), 95 – 98.
- LEVHARI, D. and SAMUELSON, P. A. (1966). The nonswitching theorem is false. *The Quarterly Journal of Economics*, **80** (4), 518–519.
- LEVY, H. and KROLL, Y. (1978). Ordering uncertain options with borrowing and lending. *Journal of Finance*, pp. 553–574.
- LUCA, M. (2011). *Reviews, reputation, and revenue: The case of Yelp. com*. Tech. rep., Harvard Business School.
- MACHINA, M. and ROTHCHILD, M. (2008). Risk. In the New Palgrave Dictionary of Economics, edited by Steven N. Durlauf and Lawrence E. Blume.
- MASSEY, D. S. and DENTON, N. A. (1988). The dimensions of residential segregation. *Social forces*, **67** (2), 281–315.
- MCVITIE, D. G. and WILSON, L. B. (1970). Stable marriage assignments for unequal sets. *BIT*, **10**, 295–309.
- and — (1971). The stable marriage problem. *Commun. ACM*, **14** (7), 486–490.
- MEYER, J. (1987). Two-moment decision models and expected utility maximization. *The American Economic Review*, pp. 421–430.

- NASH, J. (1951). Non-cooperative games. *Annals of mathematics*, **54** (2), 286–295.
- NORSTRØM, C. J. (1972). A sufficient condition for a unique nonnegative internal rate of return. *Journal of Financial and Quantitative Analysis*, **7** (03), 1835–1839.
- NUNN, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics*, **122** (2), pp. 569–600.
- PALACIOS-HUERTA, I. and VOLIJ, O. (2004). The measurement of intellectual influence. *Econometrica*, **72** (3), 963–977.
- PATHAK, P. A. and SONMEZ, T. (2008). Leveling the playing field: Sincere and sophisticated players in the boston mechanism. *The American Economic Review*, **98** (4), 1636–1652.
- PERRY, M. and RENY, P. J. (2013). *How to count citations if you must*. Tech. rep., Working paper.
- PERSICO, N. (2000). Information acquisition in auctions. *Econometrica*, **68** (1), 135–148.
- PITTEL, B. (1989). The average number of stable matchings. *SIAM J. Discret. Math.*, **2** (4), 530–549.
- (1992). On likely solutions of a stable marriage problem. *The Annals of Applied Probability*, **2** (2), pp. 358–401.
- , SHEPP, L. and VEKLEROV, E. (2007). On the number of fixed pairs in a random instance of the stable marriage problem. *SIAM J. Discrete Math.*, pp. 947–958.
- POST, T., VAN DEN ASSEM, M. J., BALTUSSEN, G. and THALER, R. H. (2008). Deal or no deal? decision making under risk in a large-payoff game show. *The American economic review*, pp. 38–71.
- PRATT, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, **32** (1/2), pp. 122–136.
- and HAMMOND, J. S. (1979). Evaluating and comparing projects: Simple detection of false alarms. *The Journal of Finance*, **34** (5), 1231–1242.

- REARDON, S. F. and FIREBAUGH, G. (2002). Measures of multigroup segregation. *Sociological methodology*, **32** (1), 33–67.
- ROTH, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of Operations Research*, **7**, 617–628.
- (1985). The college admissions problem is not equivalent to the marriage problem. *Journal of Economic Theory*, **36** (2), 277–288.
- (1990). New physicians: a natural experiment in market organization. *Science*, **250** (4987), 1524–1528.
- and PERANSON, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, **89** (4), 748–780.
- and ROTHBLUM, U. G. (1999). Truncation strategies in matching markets - in search of advice for participants. *Econometrica*, **67**, 21–43.
- and VANDE VATE, J. H. (1991). Incentives in two-sided matching with random stable mechanisms. *Economic Theory*, **1** (1), 31–44.
- and XING, X. (1994). Jumping the gun: Imperfections and institutions related to the timing of market transactions. *American Economic Review*, **84** (4), 992–1044.
- ROTHSCHILD, M. and STIGLITZ, J. E. (1970). Increasing risk: I. a definition. *Journal of Economic Theory*, **2** (3), 225–243.
- SAMUELSON, P. A. (1960). The st. petersburg paradox as a divergent double limit. *International Economic Review*, **1** (1), 31–37.
- (1963). Risk and uncertainty-a fallacy of large numbers. *Scientia*, **98** (612), 108.
- SCHREIBER, A. (2013). Economic indices of absolute and relative riskiness. *Economic Theory*, pp. 1–23.
- SHAPLEY, L. S. and SHUBIK, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, **48** (03), 787–792.
- SHARPE, W. F. (1966). Mutual fund performance. *The Journal of Business*, **39** (1), pp. 119–138.



- SHORRER, R. (2014). Consistent indices. *mimeo*.
- (2015). Entropy and the value of information for investors: the prior free implications. *mimeo*.
- SÖNMEZ, T. (1997). Manipulation via capacities in two-sided matching markets. *Journal of Economic Theory*, **77**, 197–204.
- TURVEY, R. (1963). Present value versus internal rate of return—an essay in the theory of the third best. *The Economic Journal*, **73** (289), 93–98.
- WELCH, I. (2008). *Corporate finance: an introduction*. Prentice Hall.
- WILLINGER, M. (1989). Risk aversion and the value of information. *Journal of Risk and Insurance*, pp. 104–112.
- YAARI, M. E. (1969). Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory*, **1** (3), 315–329.
- YITZHAKI, S. (1983). On an extension of the gini inequality index. *International Economic Review*, pp. 617–628.

# Appendix A

## Appendix to Chapter 1

### A.1 Proofs

#### A.1.1 Fact A.1 and Lemmata A.1 and A.2

**Definition** (Full-image). An index of riskiness  $Q$  satisfies *full-image* if for every  $\epsilon > 0$ ,  $\text{Im } Q(\mathcal{G}_\epsilon) = \mathbb{R}_+$ .

full-image says that even when the support of the gambles is limited to an  $\epsilon$ -ball, the image of  $Q$  is all of  $\mathbb{R}_+$ . Both  $Q^{AS}$  and  $Q^{FH}$  satisfy full-image. This is simply demonstrated by considering gambles of the form  $g = [\epsilon, \frac{e^{c\epsilon}}{1+e^{c\epsilon}}; -\epsilon, \frac{1}{1+e^{c\epsilon}}]$  and  $g' = [\epsilon, \frac{1}{2}; -\frac{\epsilon}{1+\epsilon c}, \frac{1}{2}]$ , as  $Q^{AS}(g) = \frac{1}{c}$  and  $Q^{FH}(g') = \frac{1}{c}$ .

**Fact A.1.** *If  $Q$  satisfies full-image then  $R_Q(u, w) \geq S_Q(u, w)$  for every  $u$  and  $w$ .*

*Proof.* By the properties of the supremum, since

$$\begin{aligned} & \{Q(g) | g \in \mathcal{G}_\epsilon \text{ and } g \text{ is accepted by } u \text{ at } w\} \\ & \cup \{Q(g) | g \in \mathcal{G}_\epsilon \text{ and } g \text{ is rejected by } u \text{ at } w\} = \mathbb{R}_+. \end{aligned}$$

If the supremum of the first set is less than the infimum of the second, then intermediate points do not belong to either in violation of full-image.  $\square$

**Lemma A.1.** *If  $Q$  satisfies homogeneity and  $0 < S_Q(u, w) < \infty$  for all  $u$  and  $w$ , then  $Q$  satisfies full-image.*

*Proof.* For some  $u$  and  $w$ ,  $S_Q(u, w) = c$ ,  $0 < c < \infty$ . Hence for some small positive  $\epsilon'$ , for every  $0 < \epsilon < \epsilon'$  there exists gambles in  $\mathcal{G}_\epsilon$  with  $Q$ -riskiness greater than  $\frac{\epsilon}{2}$ . Since multiplying by  $0 < \lambda < 1$  keeps the gambles in  $\mathcal{G}_\epsilon$ , there are gambles with any level of  $Q$ -riskiness lower than  $\frac{\epsilon}{2}$  in  $\mathcal{G}_\epsilon$ . Since for  $\lambda > 1$ ,  $\epsilon < \epsilon'$  implies that  $\frac{\epsilon}{\lambda} < \epsilon'$ , the same applies to  $\mathcal{G}_{\frac{\epsilon}{\lambda}}$ . But, using homogeneity, this means that  $\mathcal{G}_\epsilon$  includes gambles with any level of  $Q$ -riskiness lower than  $\lambda \cdot \frac{\epsilon}{2}$ . Since  $\lambda > 1$  was arbitrary, the proof is complete.  $\square$

**Lemma A.2.** *If  $Q$  satisfies homogeneity and local consistency, then  $0 < S_Q(u, w) = R_Q(u, w) < \infty$  for all  $u$  and  $w$ .*

*Proof.* Local consistency states that

$$\forall u \ \forall w \ \exists \lambda > 0 \ \forall \delta > 0 \ \exists \epsilon > 0 \ R_Q^\epsilon(u, w) - \delta < \lambda < S_Q^\epsilon(u, w) + \delta,$$

which implies that

$$\forall u \ \forall w \ \exists \lambda > 0 \ R_Q(u, w) \leq \lambda \leq S_Q(u, w).$$

Since for any  $u, w$ , and  $\epsilon > 0$  the set  $\{g \mid g \in \mathcal{G}_\epsilon, g \text{ is rejected by } u \text{ at } w\}$  is non empty, there exists a sequence of gambles  $\{g_n\}$  such that for each  $n$   $g_n$  is rejected,  $g_n \in \mathcal{G}_{\frac{1}{n}}$  and  $Q(g_n) < (1 + \frac{1}{n}) \cdot S_Q^{1/n}(u, w)$ . For small  $\delta > 0$ , let  $h_n := (1 - \delta)g_n$  for each  $n$ . For  $n$  large enough,  $h_n$  are all accepted since  $Q(h_n) = (1 - \delta)^k Q(g_n) < S_Q^{1/n}(u, w)$  and  $h_n$  is in  $\mathcal{G}_{\frac{1}{n}}$ . But this implies that  $R_Q(u, w) > (1 - \delta)^k S_Q(u, w)$  since  $h_n$  are almost always accepted and  $\lim_{n \rightarrow \infty} Q(h_n) = (1 - \delta)^k \lim_{n \rightarrow \infty} Q(g_n) = (1 - \delta)^k S_Q(u, w)$ . Since  $\delta$  was arbitrarily small, this implies  $R_Q(u, w) \geq S_Q(u, w)$ . So, putting the results together, gives

$$\forall u \forall w \exists \lambda > 0 \lambda \leq S_Q(u, w) \leq R_Q(u, w) \leq \lambda,$$

which completes the proof.  $\square$

### A.1.2 Theorem 1.2

*Proof.* (i) I first show that for every  $a > 0$  any combination of the form  $Q_a(g) := Q^{FH}(g) + a \cdot |Q^{FH}(g) - Q^{AS}(g)|$  is an index of riskiness for which the coefficient of local aversion equals the coefficient of local aversion to  $Q^{FH}$ . The reason is that for small supports, the second element in the definition is vanishingly small by Inequality 1.3.0.3, and so  $Q_a$  and  $Q^{FH}$  should be close.

Fix  $a > 0$ . First, note that

$$\forall g \in \mathcal{G} \ 0 < Q^{FH}(g) \leq Q^{FH}(g) + a \cdot |Q^{FH}(g) - Q^{AS}(g)|,$$

so  $Q_a(g) \in \mathbb{R}_+$ . Additionally, for every  $\delta > 0$  there exists  $\epsilon > 0$  small enough such that for every  $g \in \mathcal{G}_\epsilon$ ,

$$Q^{FH}(g) \leq Q^{FH}(g) + a \cdot |Q^{FH}(g) - Q^{AS}(g)| \leq Q^{FH}(g) + \delta. \quad (\text{A.1.2.1})$$

Inequality A.1.2.1 stems from the small support combined with Inequality 1.3.0.3. It tells us that the coefficient of local aversion to  $Q_a$ -riskiness cannot be different from  $A_{Q^{FH}}$  which equals  $A_{Q^{AS}}$  according to Theorem 1.1. That local consistency is satisfied follows from the same reasoning. The proof of (i) is completed by recalling that  $Q^{FH} \neq Q^{AS}$ , that both indices are locally consistent (immediate from Theorem 1.1) and homogeneous.<sup>1</sup>

(ii) Follows from Example 1.2. □

### A.1.3 Theorem 1.3

*Proof.* I start with the first part. In one direction,  $\rho_u(w) > \rho_v(w')$  implies that  $(u, w) \succ (v, w')$  (Yaari, 1969), so Lemma 1.1 implies that  $A_Q(u, w) \geq A_Q(v, w')$ .

To see that  $A_Q(u, w) \neq A_Q(v, w')$ , define  $c := \left( \frac{\rho_u(w) + \rho_v(w')}{2} \right)^{-1}$ . Let  $\{g_n\}_{n=1}^\infty$  be a sequence of gambles such that  $g_n \in \mathcal{G}_{\frac{1}{n}}$  and  $Q^{AS}(g_n) = c$ . For a small  $\delta > 0$  let  $h_n = (1 + \delta)g_n$ . By Theorem 1.1, for large values of  $n$ ,  $g_n$  and  $h_n$  will be rejected by  $u$  at  $w$  and accepted by  $v$  at  $w'$ , so

$$S_Q(v, w') \geq R_Q(v, w') \geq (1 + \delta)^k \cdot S_Q(u, w) > S_Q(u, w) \geq R_Q(u, w),$$

where the strict inequality follows from the fact that  $\infty > S_Q(u, w) > 0$  by Lemma A.2, the first and the last inequality follow from the local consistency axiom, and the second inequality follows from the definitions of  $R_Q$  and

---

<sup>1</sup>An alternative proof could use indices of the form:  $(Q^{FH})^\alpha (Q^{AS})^{1-\alpha}$ ,  $\alpha \in (0, 1)$ . This form may prove to be useful in empirical work, since it enables some flexibility in the estimation. In addition, it allows us to put some weight on the FH measure that “punishes” heavily for rare disasters (Barro, 2006).

$S_Q$  and homogeneity, by the properties of  $g_n$  and  $h_n$ . This proves that  $A_Q(u, w) > A_Q(v, w')$ .

In the other direction, if  $A_Q(u, w) > A_Q(v, w')$  then, from homogeneity and the fact that  $\infty > R_Q(v, w') > R_Q(u, w) > 0$ , there exists a sequence of gambles  $\{k_n\}_{n=1}^\infty$  such that  $k_n \in \mathcal{G}_{\frac{1}{n}}$  and  $Q(k_n) = c'$ , where  $c' := \left( \frac{A_Q(u, w) + A_Q(v, w')}{2} \right)^{-1}$ . For a small  $\delta > 0$  let  $l_n = (1 + \delta)g_n$ . A similar argument shows that

$$S_{Q^{AS}}(v, w') = R_{Q^{AS}}(v, w') \geq (1 + \delta) \cdot S_{Q^{AS}}(u, w) > S_{Q^{AS}}(u, w) = R_{Q^{AS}}(u, w),$$

where the strict inequality follows from the fact that  $S_{Q^{AS}}(u, w) > 0$  by Lemma 1.2, the equalities follow from the same lemma, and the weak inequality follows from the definitions of  $R_{Q^{AS}}$  and  $S_{Q^{AS}}$  and the homogeneity of  $Q^{AS}$ , by the properties of  $g_n$  and  $l_n$ . Using Lemma 1.2 once again, this implies that  $\rho_u(w) > \rho_v(w')$ .

For the second part, recall that  $u$  at  $w$  is at least as averse to  $Q$ -riskiness as  $v$  at  $w'$  if for every  $\delta > 0$  there exists  $\epsilon > 0$  such that  $S_Q^\epsilon(v, w') \geq R_Q^\epsilon(u, w) - \delta$ . This implies that  $S_Q(v, w') \geq R_Q(u, w)$ , which from Lemma A.2 implies that  $R_Q(v, w') \geq R_Q(u, w)$ .

In the other direction, if  $R_Q(v, w') \geq R_Q(u, w)$ , then by Lemma A.2  $\infty > S_Q(v, w') \geq R_Q(u, w) > 0$ . This means that for every  $\delta > 0$  there exists  $\epsilon > 0$  such that  $S_Q^\epsilon(v, w') \geq R_Q^\epsilon(u, w) - \delta$ , as  $S_Q$  is the limit of  $S_Q^\epsilon$  and  $R_Q$  is the limit of  $R_Q^\epsilon$ .  $\square$

#### A.1.4 Proposition 1.2

**Definition** (Wealth-independent compound gamble (Foster and Hart, 2013)).

An index  $Q$  has the *wealth-independent compound gamble* property if for every compound gamble of the form  $f = g + \mathbf{1}_A h$ , where  $Q(g) = Q(h)$ ,  $\mathbf{1}_A$  an indicator,  $A$  is an event such that  $g$  is constant on  $A$  ( $g|_A \equiv x$  for some  $x$ ) and  $h$  is independent of  $A$ ,  $Q(f) = Q(g)$ .

*Proof.* Foster and Hart (2013) show that  $Q^{AS}$  satisfies wealth-independent compound gamble. If  $Q^{AS}(g) \neq Q^{AS}(h)$ , take the one with higher (lower) level of AS riskiness, and increase (decrease) all its values by  $\epsilon$  large enough to equate the level of riskiness of the two gambles. Use wealth independent compound gamble and monotonicity with respect to stochastic dominance to deduce the required conclusion.  $\square$

#### A.1.5 Theorems 1.4 and 1.5

*Proof.* The Theorems follow from Theorems 1.6 and 1.7 by Claims 1.2 and 1.3. For a direct proof of Theorem 1.5, let  $Q$  be as in the statement. Take some CARA function,  $u$ , and an arbitrary wealth level  $w_0$ , and observe that

$$S_Q^\infty(u, w_0) \geq \inf_w S_Q(u, w) = S_Q(u, w_0) \geq R_Q(u, w_0) = \sup_w R_Q(u, w) \geq R_Q^\infty(u, w_0).$$

The equalities follow from the lack of wealth effects in CARA functions acceptance and rejection decisions, and the middle inequality follows from reflexivity.

The inequality suggests that all rejected gambles are (weakly)  $Q$ -riskier than all accepted ones. Using monotonicity, continuity, and continuity of  $u$ ,

for each accepted gamble there exists  $\epsilon > 0$  small enough such that if reduced from all the realizations of the gamble, the resulting gamble will still be accepted. Hence, the ranking is in fact strict.

Iterating the above argument with all other possible (C)ARA values proves that  $Q$  refines the order that  $Q^{AS}$  yields (recall that CARA functions accept or reject according to a  $Q^{AS}$  riskiness cutoff, which is the inverse of their ARA coefficient). Finally, continuity implies that the index must induce the same order as  $Q^{AS}$ . That  $Q^{AS}$  satisfies the properties follows from the discussion above.  $\square$

#### A.1.6 Theorem 1.7

*Proof.* First, observe that for any CARA utility function  $u$  it must be the case that  $u$  is globally at least as averse to  $Q$ -riskiness as  $u$ , by reflexivity and the lack of wealth effects in CARA functions. Now consider two gambles  $g$  and  $g'$  with  $Q^{AS}(g) > Q^{AS}(g')$ . Consider  $u$  CARA with  $\rho_u \equiv \frac{2}{Q^{AS}(g) + Q^{AS}(g')}$ .  $u$  accepts  $g'$  and rejects  $g$ , implying that  $Q(g) \geq Q(g')$ , since otherwise strong global consistency will be violated (the violation would be the fact that  $u$  is globally no less averse to  $Q$ -riskiness than itself,  $u$  accepts  $g'$  with  $Q(g') > Q(g)$ , but rejects  $g$ ).

Next, I claim that if  $Q^{AS}(g) > Q^{AS}(g')$ , but  $Q(g) = Q(g')$ , then there exists a gamble  $g_\epsilon$  such that  $Q^{AS}(g_\epsilon) > Q^{AS}(g')$ , but  $Q(g_\epsilon) < Q(g')$  in contradiction to the above result. To see this note that from monotonicity of  $Q$ , for any small  $\epsilon > 0$  a gamble  $g_\epsilon = g + \epsilon$  has  $Q(g_\epsilon) < Q(g)$ , and from continuity of  $Q^{AS}$ , for small enough  $\epsilon$ ,  $Q^{AS}(g_\epsilon) > Q^{AS}(g')$ .



Finally, I claim that if  $Q^{AS}(g) = Q^{AS}(g')$ , but  $Q(g) > Q(g')$ , then there exists a gamble  $g_\epsilon$  such that  $Q^{AS}(g_\epsilon) < Q^{AS}(g')$ , but  $Q(g_\epsilon) > Q(g')$ . To see this, apply the same argument from the previous paragraph, only this time use the continuity of  $Q$  and the monotonicity of  $Q^{AS}$ .

The upshot of the above discussion is that  $Q^{AS}(g) > Q^{AS}(g') \iff Q(g) > Q(g')$  as required.

□

### A.1.7 Theorem 1.8

**Lemma A.3.** *If  $Q$  is a continuous index of performance that satisfies global consistency, reflexivity, translation invariance, monotonicity and homogeneity, and  $u$  and  $v$  are two CARA utilities with  $\rho_u \leq \rho_v$ , then  $u$  is globally inclined to invest in  $Q$ -performers at least as  $v$ .*

*Proof.* From reflexivity and the fact that there are no wealth effects for CARA functions it follows that  $u$  is globally inclined to invest in  $Q$ -performers at least as itself. The conclusion follows, as for any  $w, w', v$  accepts less transactions at  $w'$  than  $u$  at  $w$  in the sense of set inclusion, so for all  $\bar{q} > q > 0$  and  $\delta > 0$ , there exists  $\epsilon > 0$  such that

$$\begin{aligned} 0 &\leq \sup_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is rejected by } u \text{ at } w\} \\ &\leq \inf_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is accepted by } u \text{ at } w'\} + \delta \\ &\leq \inf_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is accepted by } v \text{ at } w'\} + \delta, \end{aligned}$$

where  $\bar{q}$  is the value that is used for reflexivity at  $(u, w)$ .

□

**Lemma A.4.** *The following are equivalent:*

- (i)  *$u$  at  $w$  is locally inclined to invest in  $P^{AS}$ -performers at least as  $v$  at  $w'$*
- (ii)  $\rho_u(w) \leq \rho_v(w')$

*Proof.*  $\neg(ii) \implies \neg(i)$ : By Theorem 1.1 if  $\rho_u(w) > \rho_v(w')$ , then for small enough  $\epsilon > 0$   $v$  at  $w'$  accepts any local transaction such that  $Q^{AS}(q \cdot r) = \frac{3}{2\rho_u(w) + \rho_v(w')}$  or  $Q^{AS}(q \cdot r) = \frac{3}{\rho_u(w) + 2\rho_v(w')}$ , and such transactions are rejected by  $u$  at  $w$ . Such transactions have  $P^{AS}(r) = q \cdot \frac{2\rho_u(w) + \rho_v(w')}{3}$  and  $P^{AS}(r) = q \cdot \frac{\rho_u(w) + 2\rho_v(w')}{3}$  respectively. This implies that

$$\sup_{(q,r) \in \mathcal{T}_\epsilon} \{P^{AS}(r) \mid (q,r) \text{ is rejected by } u \text{ at } w\} \geq q \cdot \frac{2\rho_u(w) + \rho_v(w')}{3} >$$

$$q \cdot \frac{\rho_u(w) + 2\rho_v(w')}{3} \geq \inf_{(q,r) \in \mathcal{T}_\epsilon} \{Q(r) \mid (q,r) \text{ is accepted by } v \text{ at } w'\}$$

, for all  $\epsilon > 0$ , so (i) does not hold (use  $\delta = \frac{\bar{q}}{3} \cdot \left| \frac{2\rho_u(w) + \rho_v(w')}{3} - \frac{\rho_u(w) + 2\rho_v(w')}{3} \right|$  to get a contradiction).

(ii)  $\implies$  (i): By Theorem 1.1 and an argument as above,  $P^{AS}$  satisfies reflexivity. Thus, for some  $\bar{q}_1$ , for all  $\bar{q}_1 > q > 0$  and all  $\delta > 0$  there exists  $\epsilon > 0$  with

$$\begin{aligned} 0 &\leq \sup_{(q,r) \in \mathcal{T}_\epsilon} \{P^{AS}(r) \mid (q,r) \text{ is rejected by } u \text{ at } w\} \leq \\ &\leq \inf_{(q,r) \in \mathcal{T}_\epsilon} \{P^{AS}(r) \mid (q,r) \text{ is accepted by } u \text{ at } w\} + \delta. \end{aligned}$$

By the same theorem, there exists  $\bar{q}_2$  such that for all  $\bar{q}_2 > q > 0$  and  $\delta$  there exists  $\epsilon' > 0$  with

$$\inf_{(q,r) \in \mathcal{T}_{\epsilon'}} \{Q(r) \mid (q,r) \text{ is accepted by } u \text{ at } w\} \leq$$

$$\leq \inf_{(q,r) \in \mathcal{T}_{\epsilon'}} \{Q(r) \mid (q,r) \text{ is accepted by } v \text{ at } w'\} + \delta.$$

Thus, for all  $\min\{\bar{q}_1, \bar{q}_2\} > q > 0$  and all  $\delta' (= 2\delta) > 0$ , there exists  $\min\{\epsilon, \epsilon'\} > \bar{\epsilon} > 0$  such that

$$\begin{aligned} 0 &\leq \sup_{(q,r) \in \mathcal{T}_{\bar{\epsilon}}} \{Q(r) \mid (q,r) \text{ is rejected by } u \text{ at } w\} \leq \\ &\leq \inf_{(q,r) \in \mathcal{T}_{\bar{\epsilon}}} \{Q(r) \mid (q,r) \text{ is accepted by } v \text{ at } w'\} + \delta'. \end{aligned}$$

□

**Lemma A.5.**  $P^{AS}$  is a continuous index of performance that satisfies reflexivity, global consistency, translation invariance, monotonicity and homogeneity.

*Proof.* Translation invariance is immediate as the index could be expressed as a function of  $r - r_f$ . From now on, assume without loss of generality that  $r_f = 0$ . For homogeneity, note that both the expectation operator and  $Q^{AS}$  are homogeneous of degree 1, and so their ratio is homogeneous of degree 0. Continuity follows from the continuity of  $Q^{AS}$  and the fact that if  $r^n$  are bounded and converge to  $r$ ,  $\mathbb{E}[r_n]$  converges to  $\mathbb{E}[r]$  from the bounded convergence theorem.

For any  $r$  if  $\mathbb{E}[r] = c > 0$  then for all  $\lambda > 0$   $\mathbb{E}[(r + \lambda)] = c + \lambda \equiv (1 + \epsilon) \mathbb{E}[r]$  for some  $\epsilon > 0$ . From homogeneity of degree 1 and monotonicity with respect to first order stochastic dominance of  $Q^{AS}$  one has

$$\begin{aligned} Q^{AS}\left(\frac{r}{\mathbb{E}[r]}\right) &= Q^{AS}\left(\frac{(1 + \epsilon)r}{\mathbb{E}[(1 + \epsilon)r]}\right) = Q^{AS}\left(\frac{(1 + \epsilon)r}{c + \lambda}\right) > \\ &> Q^{AS}\left(\frac{r}{c + \lambda}\right) > Q^{AS}\left(\frac{(r + \lambda)}{c + \lambda}\right), \end{aligned}$$

where the inequalities follows from monotonicity of  $Q^{AS}$  with respect to first order stochastic dominance, and from the homogeneity of degree 1 of  $Q^{AS}$ . The previous inequality implies that  $P_{r_f}^{AS}(r + \lambda) > P_{r_f}^{AS}(r)$ .

Reflexivity was proved in Lemma A.4. Global consistency is implied by the global consistency of  $Q^{AS}$ , by Lemma A.4.  $\square$

**Lemma A.6.** *If  $P$  is a continuous index of performance that satisfies reflexivity, global consistency, translation invariance, monotonicity and homogeneity of degree 0, then it is ordinally equivalent to  $P^{AS}$ .*

*Proof.* Assume, by way of contradiction that  $P$  satisfies the conditions but is not ordinally equivalent to  $P^{AS}$ . There are three ways such violation happen:

1. There exist  $r, r' \in \mathcal{R}_1$  with  $P^{AS}(r) > P^{AS}(r')$  and  $P(r) < P(r')$
2. There exist  $r, r' \in \mathcal{R}_1$  with  $P^{AS}(r) > P^{AS}(r')$  and  $P(r) = P(r')$
3. There exist  $r, r' \in \mathcal{R}_1$  with  $P^{AS}(r) = P^{AS}(r')$  and  $P(r) < P(r')$

There is no loss of generality in treating only the first case. The reason is that using monotonicity and continuity, we could slightly shift  $r$  and  $r'$  to break the equalities in the right direction while not effecting the inequalities.

Given a violation of type 1, consider an agent with CARA utility function,  $u$  such that  $\rho_u \equiv .6P^{AS}(r) + .4P^{AS}(r')$ . Note that  $u$  accepts  $(1, r)$  but rejects  $(1, r')$ , and that  $u$  is globally inclined to invest in  $P$ -performers at least as  $u$  by Lemma A.3. But this means that global consistency is violated by  $P$ .  $\square$

*Proof.* (Of the theorem) Follows from the lemmata.  $\square$

#### A.1.8 Lemma 1.4

**Lemma A.7.**  $g \in \mathcal{H} \iff \log(1 + g) \in \mathcal{G}$ .

*Proof.* In one direction,  $g \in \mathcal{H} \Rightarrow g \in \mathcal{G}$  and  $Q^{FH}(g) < 1$ . Since  $Q^{FH}(g) \geq L(g)$  it follows that  $\log(1+g)$  is well-defined. As  $g \in \mathcal{G}$ , it assumes a negative value with positive probability and therefore so does  $\log(1+g)$ . Finally,  $Q^{FH}(g) < 1$  implies that  $\mathbb{E}[\log(1+g)] > 0$ . Hence,  $\log(1+g) \in \mathcal{G}$ .

In the other direction, if  $\log(1+g) \in \mathcal{G}$  we have that  $\log(1+g)$  assumes a negative value with positive probability and therefore so does  $g$ . In addition, we have  $\sum p_i \log(1+g_i) > 0$ . Hence, by Fact 1.1,  $g \in \mathcal{H}$ .  $\square$

*Proof.* (of Lemma 1.4) Note that for every  $g \in \mathcal{H}$  and  $S > 0$ , we have  $\mathbb{E}\left[(1+g)^{-\frac{1}{S}}\right] = \mathbb{E}\left[e^{-\frac{\log(1+g)}{S}}\right]$ . Consequentially, Lemma A.7 and Theorem A in AS imply that the unique positive solution for the equation is  $S(g) = Q^{AS}(\log(1+g))$ .  $\square$

### A.1.9 Theorem 1.10

**Lemma A.8.** *For all  $g \in \mathcal{H}$ , If  $u \in \mathcal{U}$  has a constant RRA then  $\varrho_u(w) - 1 < \frac{1}{S(g)}$  if and only if  $\mathbb{E}[u(w+wg)] > u(w) \ \forall w > 0$ .*

*Proof.* As positive affine transformations of the utility function do not change acceptance and rejection, it is enough to treat functions of the form  $u(w) = -w^{1-\alpha}$ . Now observe that:

$$\begin{aligned} \mathbb{E}[u(w+wg)] > u(w) &\iff \mathbb{E}[-w^{1-\alpha}(1+g)^{1-\alpha}] > -w^{1-\alpha} \iff \\ &\iff \mathbb{E}[(1+g)^{1-\alpha}] < 1 \iff \mathbb{E}\left[e^{(1-\alpha)\log(1+g)}\right] < 1 \iff \\ &\iff Q^{AS}(\log(1+g)) < \frac{1}{\alpha-1} \iff \alpha-1 < \frac{1}{S(g)}. \end{aligned}$$

$\square$

**Lemma A.9.** For every  $u, v \in \mathcal{U}$ , if  $\inf_x \varrho_u(x) \geq \sup_{x'} \varrho_v(x')$  then for every  $w$ , if  $u$  accepts  $g$  at  $w$  so does  $v$ .

*Proof.* Without loss of generality, assume that  $v(w) = u(w) = 0$  and that  $v'(w) = u'(w) = 1$ . For every  $t > 1$

$$\begin{aligned} \log v'(tw) &= \log v'(tw) - \log v'(w) = \int_1^t \frac{\partial \log v'(sw)}{\partial s} ds = \int_1^t w \frac{v''(sw)}{v'(sw)} ds = \\ &= \int_1^t \frac{1}{s} \cdot \left( sw \frac{v''(sw)}{v'(sw)} \right) ds \geq \int_1^t \frac{1}{s} \cdot \left( sw \frac{u''(sw)}{u'(sw)} \right) ds = \log u'(tw) \end{aligned}$$

$$\begin{aligned} \log v'\left(\frac{w}{t}\right) &= \log v'\left(\frac{w}{t}\right) - \log v'(w) = \int_1^t \frac{\partial \log v'\left(\frac{w}{s}\right)}{\partial s} ds = \int_1^t -\frac{w}{s^2} \frac{v''\left(\frac{w}{s}\right)}{v'\left(\frac{w}{s}\right)} ds = \\ &= \int_1^t \frac{1}{s} \cdot \left( -\frac{w}{s} \frac{v''\left(\frac{w}{s}\right)}{v'\left(\frac{w}{s}\right)} \right) ds \leq \int_1^t \frac{1}{s} \cdot \left( -\frac{w}{s} \frac{u''\left(\frac{w}{s}\right)}{u'\left(\frac{w}{s}\right)} \right) ds = \log u'\left(\frac{w}{t}\right) \end{aligned}$$

This means that for every  $t > 0$ :

$$v(tw) = v(tw) - v(w) = \int_1^t w v'(sw) ds \geq \int_1^t w u'(sw) ds = u(tw)$$

And so, if  $\mathbb{E}[u(w + wg)] > u(w) = 0$  then necessarily  $\mathbb{E}[v(w + wg)] > v(w) = 0$  as  $\mathbb{E}[v(w + wg)] \geq \mathbb{E}[u(w + wg)]$ .  $\square$

**Lemma A.10.** For every  $u \in \mathcal{U}$  and every  $w > 0$ ,  $R_S(u, w) = S_S(u, w)$  and  $A_S(u, w) = \varrho_u(w) - 1$ .

The proof of Lemma A.10 is analogous to the proof of Lemma 1.2 and is

therefore omitted. Recalling that the CRRA utility function with parameter  $\alpha$  is often expressed as

$$-w^{1-\alpha} = -w^{-(\alpha-1)},$$

this transformation of  $\varrho_u(\cdot)$  seems particularly natural.

*Proof.* (Of the theorem, sketch). First observe that for every  $\alpha > 0$

$$S((1+g)^\alpha - 1) = Q^{AS}(\log(1+g)^\alpha) = Q^{AS}(\alpha \cdot \log(1+g)) = \alpha \cdot Q^{AS}(\log(1+g)) = \alpha \cdot S(g),$$

so  $S$  satisfies Scaling. By Lemma A.10,  $\infty > R_S(u, w) = S_S(u, w) = \frac{1}{\varrho_u(w)-1} > 0$  (which implies that  $S$  satisfies local consistency).

To see that  $S$  satisfies global consistency, observe that the fact that  $A_S$  is ordinally equivalent to  $\varrho$  implies that if  $v > u$  then there exist  $\lambda \geq 1$  with  $\inf_w \varrho_v(w) \geq \lambda \geq \sup_{w'} \varrho_u(w')$ . Therefore, by Lemma A.9 if  $v$  accepts  $g$  at  $w$  so does an agent with a CRRA utility function with RRA equals  $\lambda$ . Furthermore, by Lemma A.8, if  $S(h) < S(g)$  this agent will accept  $h$  at any wealth level. Applying Lemma A.9 again implies that  $u$  accepts  $h$  at  $w$ .

For uniqueness, assume that  $\hat{Q}$  satisfies the requirements. By Lemma A.7  $\hat{P}(g) := \hat{Q}(e^g - 1)$  is an index of riskiness  $\hat{P} : \mathcal{G} \rightarrow \mathbb{R}_+$ . For every  $\alpha > 0$ , we have  $\hat{P}(\alpha g) = \hat{Q}(e^{\alpha g} - 1) = \hat{Q}((1 + e^g - 1)^\alpha - 1) = \alpha \cdot \hat{Q}(e^g - 1) = \alpha \cdot \hat{P}(g)$ , so  $\hat{P}$  satisfies homogeneity. I next claim that  $\hat{Q}(g) > \hat{Q}(h)$  if and only if  $S(g) > S(h)$ . To see this, note that from Theorem 1.9  $A_{\hat{Q}}$  is ordinally equivalent to  $\varrho$  and that from local consistency and scaling  $0 < S_Q(u, w) = R_Q(u, w) < \infty$  (see Lemma A.2 for a proof of the analogous case). From these facts it follows that  $S$  and  $\hat{Q}$  order lotteries in the same manner (as before, using CRRA

functions). Hence,  $\hat{P}$  and  $Q^{AS}$  also agree on the order of lotteries. Since both  $\hat{P}$  and  $Q^{AS}$  are homogeneous, we have that  $\hat{P} = C \cdot Q^{AS}$  for some  $C > 0$ . This in turn, implies that  $\hat{Q} = C \cdot S$ , for some  $C > 0$ .  $\square$

#### A.1.10 Theorem 1.11

*Proof.* First, observe that for any CRRA utility function  $u$  it must be the case that  $u$  is globally at least as averse to  $Q$ -riskiness as  $u$ , by reflexivity and the lack of wealth effects in CRRA functions. Now consider two gambles  $g$  and  $g'$  with  $S(g) > S(g')$ . Consider  $u$  CRRA with  $\rho_u \equiv 1 + \frac{2}{S(g)+S(g')}$ .  $u$  accepts  $g'$  and rejects  $g$ , implying that  $Q(g) \geq Q(g')$ , since otherwise strong global consistency will be violated (the violation would be the fact that  $u$  is globally no less averse to  $Q$ -riskiness than itself,  $u$  accepts  $g'$  with  $Q(g') > Q(g)$ , but rejects  $g$ ).

Next, I claim that if  $S(g) > S(g')$ , but  $Q(g) = Q(g')$ , then there exists a gamble  $g_\epsilon$  such that  $S(g_\epsilon) > S(g')$ , but  $Q(g_\epsilon) < Q(g')$  in contradiction to the above result. To see this note that from monotonicity of  $Q$ , for any small  $\epsilon > 0$  a gamble  $g_\epsilon = g + \epsilon$  has  $Q(g_\epsilon) < Q(g)$ , and from continuity of  $S$ , for small enough  $\epsilon$ ,  $S(g_\epsilon) > S(g')$ .

Finally, I claim that if  $S(g) = S(g')$ , but  $Q(g) > Q(g')$ , then there exists a gamble  $g_\epsilon$  such that  $S(g_\epsilon) < S(g')$ , but  $Q(g_\epsilon) > Q(g')$ . To see this, apply the same argument from the previous paragraph, only this time use the continuity of  $Q$  and the monotonicity of  $S$ .

The upshot of the above discussion is that  $S(g) > S(g') \iff Q(g) > Q(g')$  as required.  $\square$



### A.1.11 Theorem 1.14

**Lemma A.11.** *Let  $c = (x_n, t_n)_{n=1}^N$  be an investment cashflow. If  $r_k(s) < r_j(s)$  for all  $s \in [t_1, t_N]$  then, for all  $t$ ,  $\sum_n e^{-\int_t^{t_n} r_k(s) ds} x_n \leq 0$  implies that  $\sum_n e^{-\int_t^{t_n} r_j(s) ds} x_n < 0$ .*

*Proof.* Denote by  $n^*$  the highest index with  $x_n < 0$ . Then

$$\begin{aligned} \sum_n e^{-\int_t^{t_n} r_k(s) ds} x_n &= \sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} x_n + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} x_n = \\ &= -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n|, \end{aligned} \quad (\text{A.1.11.1})$$

and

$$\begin{aligned} -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| \leq 0 &\iff \\ \iff e^{\int_t^{t_{n^*}} r_k(s) ds} \cdot \left( -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| \right) &\leq 0, \end{aligned} \quad (\text{A.1.11.2})$$

and similar statements hold when  $r_k$  is replaced with  $r_j$ . But,

$$\begin{aligned} e^{\int_t^{t_{n^*}} r_k(s) ds} \cdot \left( -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| \right) &= \\ -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_k(s) ds} |x_n| &> \\ -\sum_{n \leq n^*} e^{-\int_t^{t_n} r_j(s) ds} |x_n| + \sum_{n > n^*} e^{-\int_t^{t_n} r_j(s) ds} |x_n| & \end{aligned}$$

as positives are only multiplied by smaller numbers and negatives are multiplied by greater (positive) numbers.  $\square$

**Lemma A.12.** *If  $c = (x_n, t_n)_{n=1}^N$  is an investment cashflow then there exists a unique positive number  $r$  such that  $\sum_n e^{-rt_n} x_n = 0$ . Furthermore, if  $\tilde{r}(t) > r > \hat{r}(t)$  for all  $t \in [t_1, t_N]$ , then the NPV of  $c$  is negative using  $\tilde{r}$ , and is positive using  $\hat{r}$ .*

For general cashflows, multiple solutions to the equation defining the internal rate of return may exist. Interestingly, both Arrow and Pratt took interest in finding simple conditions that would rule out this possibility (Arrow and Levhari, 1969; Pratt and Hammond, 1979). Lemma A.12 generalizes the result of Norström (1972) who had shown that investment cashflows have a unique positive IRR in the discrete setting.

*Proof.* Define the function  $f(\alpha) := \sum_n e^{-\alpha t_n} x_n$ . Observe that  $f(\cdot)$  is continuous, and satisfies  $f(0) > 0$  and  $f(\alpha) < 0$  for large values of  $\alpha$ . Hence, continuity implies the existence of a solution. Lemma A.11 implies its uniqueness, and the second part of the claim.  $\square$

**Lemma A.13.** *If  $T$  satisfies translation invariance, homogeneity and local consistency, then for all  $u, w$ ,  $0 < S_T(i, t) = R_T(i, t) < \infty$ .*

*Proof.* Local consistency requires that

$$\forall i \ \forall t \ \exists \lambda > 0 \ \forall \delta > 0 \ \exists \epsilon > 0 \ R_T^\epsilon(i, t) - \delta < \lambda < S_T^\epsilon(i, t) + \delta,$$

which implies that

$$\forall i \ \forall t \ \exists \lambda > 0 \ R_T(i, t) \leq \lambda \leq S_T(i, t).$$

Since for any  $i, t$ , and  $\epsilon > 0$  the set  $\{c \mid c \in \mathcal{C}_{t, \epsilon}, c \text{ is rejected by } i\}$  is non-

empty, there exists a sequence of cashflows  $\{c_n\}$  such that for each  $n$ ,  $c_n := (x_i^n, t_i^n)$  is rejected,  $c_n \in \mathcal{C}_{t, \frac{1}{n}}$  and  $T(c_n) < (1 + \frac{1}{n}) \cdot S_T^{1/n}(i, t)$ . For small  $\delta > 0$ , let  $c'_n := (x_i^n, (t_i - t_1)(1 - \delta))$  for each  $n$ . For  $n$  large enough,  $c'_n$  are all accepted since  $T(c'_n) = (1 - \delta)^k T(c_n) < S_T^{1/n}(i, t)$  and  $c'_n$  is in  $\mathcal{C}_{t, \frac{1}{n}}$ . But this implies that  $R_T(i, t) > (1 - \delta)^k S_T(i, t)$  since  $c'_n$  are almost always accepted and  $\lim_{n \rightarrow \infty} T(c'_n) = (1 - \delta)^k \lim_{n \rightarrow \infty} T(c_n) = (1 - \delta)^k S_Q(i, t)$ . Since  $\delta$  was arbitrarily small, this implies  $R_T(i, t) \geq S_T(i, t)$ . So, putting the results together, gives

$$\forall i \ \forall t \ \exists \lambda > 0 \ \lambda \leq S_T(i, t) \leq R_T(i, t) \leq \lambda,$$

which completes the proof.  $\square$

*Proof.* (of the theorem) For the first part, in one direction, if  $r_i(t) > r_j(t')$  then there exists a small  $\epsilon' > 0$  such that for all  $x, y \in (-\epsilon', \epsilon')$   $r_i(t+x) > r_j(t+y)$ .<sup>2</sup> For a sequence of cashflows with small support and IRR of  $\frac{r_i(t)+r_j(t')}{2}$  their translations which start at  $t'$  are almost always accepted, and the translations which starts at  $t$  are almost always rejected. The same applies to these translated cashflows with times  $t_i^n$  replaced by  $(1-\delta)(t_i^n - t)$ . By Lemma A.13, homogeneity and translation invariance this implies that  $R_T(i, t) < R_T(j, t')$ .

In the other direction, assume  $R_T(i, t) < R_T(j, t')$ . From Lemma A.13  $0 < R_T(i, t) < R_T(j, t') < \infty$ . Consider a sequence of cashflows  $\{c_n\}$  with  $t_N^n < \frac{1}{n}$ ,  $t_1^n = 0$  and  $T(c_n) = \frac{2R_T(i, t) + R_T(j, t')}{3}$ . For small  $\delta$ , let  $\{c'_n\}$  be a sequence of cashflows such that  $t_i'^n = t_i^n \cdot (1 - \delta)$ . The translations of both  $\{c_n\}$  and  $\{c'_n\}$  which start at  $t'$  are almost always accepted by  $j$  and both

---

<sup>2</sup>The proof follows closely the proof of Theorem 1.3, which provides more details.

the translations that start at  $t$  are almost always rejected by  $i$ . This, in turn, implies that  $r_i(t) > r_j(t')$  using the previous Lemmata.

The second part follows from the first part and from Lemma A.13.  $\square$

### A.1.12 Propositions 1.10 and 1.11

*Proof.* (Proposition 1.10) Note that  $\forall i, t \ A_D(i, t) = r_i(t)$ . The conclusion follows from Lemma A.12.  $\square$

*Proof.* (Proposition 1.11) Follows from Lemma A.11.  $\square$

### A.1.13 Theorem 1.15

*Proof.* To prove (i) I first identify one such index. The construction draws upon the findings of previous sections. First, denote by  $\mathcal{C}^1$  the class of investment cashflows with  $|t_N - t_1| = 1$ . Restricting attention to this class of cashflows, I define a function from  $\mathcal{C}^1$  to  $\mathcal{G}$ , the class of gambles,  $\mathcal{T} : \mathcal{C}^1 \rightarrow \mathcal{G}$ ,

$$\mathcal{T}(c) = \left[ 1, \frac{e^{\frac{1}{D(c)}}}{1 + e^{\frac{1}{D(c)}}}; -1, \frac{1}{1 + e^{\frac{1}{D(c)}}} \right].$$

Observe that  $Q^{AS}(\mathcal{T}(\cdot)) \equiv D(\cdot)$ . Now, given a cashflow  $c = (x_n, t_n)_{n=1}^N$ , let  $\alpha_c := |t_N - t_1|$ . Given  $t$ , define  $\hat{c}_t := \left( x_n, t + \frac{1}{\alpha_c} (t_n - t) \right)_{n=1}^N$ . By construction,  $\hat{c}_t$  is a member of  $\mathcal{C}^1$ . This allows defining a new index  $Z : \mathcal{C} \rightarrow \mathbb{R}_+$  in the following way:

$$Z(c) := Q^{FH}(\alpha_c \cdot \mathcal{T}(\hat{c}_t)).$$

$Z$  is homogeneous and translation invariant since  $Q^{FH}$  is homogeneous, and  $\mathcal{T}$  was constructed to assure these properties.

Noting that for  $c \in \mathcal{C}_{t,\epsilon}$

$$|D(c) - Z(c)| = |Q^{AS}(\alpha_c \cdot \mathcal{T}(\hat{c}_t)) - Q^{FH}(\alpha_c \cdot \mathcal{T}(\hat{c}_t))| \leq 2\alpha_c \leq 2\epsilon,$$

one observes that  $R_Z(\cdot, \cdot) = R_D(\cdot, \cdot)$  and  $S_Z(\cdot, \cdot) = S_D(\cdot, \cdot)$ , so if  $D$  is locally consistent so is  $Z$ .

$D$  satisfies all the requirements of the theorem (proved later on) and the coefficient of local aversion to  $D$  equals to  $r$ . Since the relation at least as averse to  $D$ -delay induces the same order as  $r$ , the same applies to  $Z$ -delay, as  $0 < A_D < \infty$ . This implies that for  $a > 0$  combinations of the form  $W_a(\cdot) = Z(\cdot) + a|D(\cdot) - Z(\cdot)|$  also satisfy the requirements of (i). To see that  $D \neq Z$ , it is enough to consider a cashflow  $c$  with  $\alpha_c = 1$  and  $D(c) = 1$ . For this cashflow  $Z(c) \approx 1.26$ . Together with the fact that  $Z$  and  $D$  are uniformly close on small domains, the fact that the coefficient of local aversion to  $Z$  equals to  $r$  (which is positive and finite) implies that the same holds for  $W_a$ , which completes the proof of this part.

(ii) Follows from example A.1. □

**Example A.1.** Consider  $W_1(\cdot)$  and a cashflow  $c$  with  $\alpha_c = 1$  for which  $D(c) = 1$ . This implies that  $Z(c) \approx 1.26$ , hence  $W_1(c) < 1.6$ . Now consider another cashflow,  $c'$ , with  $\alpha_{c'} = 1$ , which first order time dominates  $c$  and has  $D(c') = \epsilon$  for a small  $\epsilon$ .<sup>3</sup> Since  $Z(c) \geq 1$  from the properties of  $Q^{FH}$  and  $\mathcal{T}$ ,  $W_1(c') > 1.6$ . Therefore, while  $c'$  first order time dominates  $c$ ,  $W_1(c) < W_1(c')$ .

---

<sup>3</sup>This could be achieved by increasing  $x_N$ .

### A.1.14 Theorem 1.16

*Proof.* I provide the proof for the case  $k = 1$ , but the generalization is simple.

First, I check that  $D$  satisfies the axioms. Homogeneity is clearly satisfied as

$$\begin{aligned} \sum_n e^{-rt_n} x_n = 0 &\iff e^{rt} \sum_n e^{-rt_n} x_n = 0 \iff \\ \sum_n e^{-r(t_n-t)} x_n = 0 &\iff \sum_n e^{-\frac{r}{\lambda} \cdot \lambda(t_n-t)} x_n = 0 \quad (\forall t \ \forall \lambda > 0). \end{aligned}$$

Translation invariance is also satisfied as

$$\sum_n e^{-rt_n} x_n = 0 \iff e^{rt} \sum_n e^{-rt_n} x_n = 0 \quad (\forall t).$$

For local consistency, I use the smoothness of  $r_i(\cdot)$  to deduce that for every  $t$  and small  $\epsilon > 0$  there exists  $\delta > 0$  such that if  $s \in (t - \delta, t + \delta)$  then  $r_i(t) - \epsilon < r_i(s) < r_i(t) + \epsilon$ . This fact, together with Lemmata A.11 and A.12, implies that  $0 < S_D(i, t) = R_D(i, t) < \infty$  and that  $A_D(i, t) = r_i(t)$ , hence the axiom is satisfied.

To see that global consistency is satisfied, first note that  $i$  is at least as averse to  $D$ -delay as  $j$  if and only if  $\sup_t r_j(t) \leq \inf_t r_i(t)$ . Consider an agent that discounts at the constant rate  $\nu$ , with  $\sup_t r_j(t) \leq \nu \leq \inf_t r_i(t)$ . Label this agent  $\nu$ . Lemma A.11 implies that  $\nu$  accepts any cashflow accepted by  $i$ , Lemma A.12 implies that he also accepts cashflows with higher IRR, and another application of Lemma A.11 implies that  $j$  accepts these cashflows.

I now turn to show that the only indices that satisfy the five axioms are positive multiples of  $D$ . This is done in two steps. In the first step, I show that indices that satisfy the axioms agree with the order induced by  $D$ . Then, I show that they are also multiples of this index.

For the first step, assume by way of contradiction that there exists another index,  $Q$ , that satisfies the axioms but does not agree with  $D$  on the ordering of two cashflows at some given time points. There are three possibilities:

1.  $Q(c) > Q(c')$  and  $D(c) < D(c')$  for cashflows  $c$  and  $c'$ .
2.  $Q(c) > Q(c')$  and  $D(c) = D(c')$  for cashflows  $c$  and  $c'$ .
3.  $Q(c) = Q(c')$  and  $D(c) < D(c')$  for cashflows  $c$  and  $c'$ .

There is no loss of generality in treating just the first case. To see this, note that the second and third cases imply the existence of an example of the first type. Such example is obtained by breaking the tie in the correct direction, using translation invariance and homogeneity, while preserving the strict inequality.

To obtain a contradiction, choose  $r_1$  and  $r_2$  such that

$$D(c) < \frac{1}{r_2} < \frac{1}{r_1} < D(c'),$$

and consider two agents that discount with the constant rates  $r_1$  and  $r_2$ , and are labeled accordingly  $r_1$  and  $r_2$  (with a slight abuse of notation). Using Lemma A.12 both  $r_1$  and  $r_2$  accept  $c$  and reject  $c'$ . Theorem 1.14 and Lemma A.13 imply that  $r_1 \prec_Q r_2$ . But this means that  $Q$  violates global consistency, as  $r_2$ , the impatient agent, accepts  $c$ , the  $Q$ -delayed cashflow, but  $r_1$  does not accept  $c'$  which is less  $Q$ -delayed. Thus,  $Q$  and  $D$  must agree on the ordering of any two cashflows at any given time point.

For the second step, choose an arbitrary cashflow  $c_0 = (x_n, t_n)_{n=1}^N$  and an index that satisfies the axioms,  $T$ . For any cashflow  $c$ , there exists a positive number  $\lambda > 0$  such that  $T\left((x_n, t_1 + \lambda \cdot (t_n - t_1))_{n=1}^N\right) = T(c)$ . The first step

implies that  $D\left((x_n, t_1 + \lambda \cdot (t_n - t_1))_{n=1}^N\right) = D(c)$ . But  $D\left((x_n, t_1 + \lambda \cdot (t_n - t_1))_{n=1}^N\right) = \lambda \cdot D(c_0)$ , and also  $T\left((x_n, t_1 + \lambda \cdot (t_n - t_1))_{n=1}^N\right) = \lambda \cdot T(c_0)$ . Altogether this means that  $T(c) = \frac{T(c_0)}{D(c_0)} D(c)$  for every  $c$ .  $\square$

#### A.1.15 Theorem 1.17

*Proof.* First, observe that for any agent with constant discount rate, it must be the case that the agent is globally at least as averse to  $T$ -delay as himself, by reflexivity and the invariance of the sign of the NPV of translations of a cashflow when the discount rate is constant. Now consider two cashflows  $c$  and  $c'$  with  $D(c) > D(c')$ . Consider  $i$  with  $r_i \equiv \frac{2}{D(c)+D(c')}$ .  $i$  accepts  $c'$  and rejects  $c$ , implying that  $T(c) \geq T(c')$ , since otherwise strong global consistency will be violated (the violation would be the fact that  $i$  is globally at least as averse to  $T$ -delay as itself,  $i$  accepts  $c'$  with  $T(c') > T(c)$ , but rejects  $c$ ).

Next, I claim that if  $D(c) > D(c')$ , but  $T(c) = T(c')$ , then there exists a cashflow  $c_\epsilon$  such that  $D(c_\epsilon) > D(c')$ , but  $T(c_\epsilon) < T(c')$  in contradiction to the above result. To see this note that from monotonicity of  $T$ , for any small  $\epsilon > 0$ , given  $c = (x_i, t_i)_{i=1}^N$ , a cashflow  $c_\epsilon = (x_i + \epsilon, t_i)_{i=1}^N$  has  $T(c_\epsilon) < T(c)$ , and from continuity of  $D$ , for small enough  $\epsilon$ ,  $D(c_\epsilon) > D(c')$ .

Finally, I claim that if  $D(c) = D(c')$ , but  $T(c) > T(c')$ , then there exists a cashflow  $c_\epsilon$  such that  $D(c_\epsilon) < D(c')$ , but  $T(c_\epsilon) > T(c')$ . To see this, apply the same argument from the previous paragraph, only this time use the continuity of  $T$  and the monotonicity of  $D$ .

The upshot of the above discussion is that  $D(c) > D(c') \iff T(c) > T(c')$  as required.  $\square$



### A.1.16 Theorem 1.20

*Proof.* (i) The proof is similar to the proof of Theorem 1.1. First, note that if  $\left\{a_n = (\mu_n, \alpha_n) \in \mathcal{A}_{\frac{1}{n}}\right\}_{n=1}^{\infty}$  are accepted it must be the case that  $\mu_n \xrightarrow{n \rightarrow \infty} 0$ . To see this, assume by way of contradiction that there is a sub-sequence of such transactions where the price does not converge to 0, without loss of generality  $a_n = (\mu_n, \alpha_n)$ , and  $\lim_{n \rightarrow \infty} \mu_n = \hat{\mu} \in (0, \infty]$ . Let  $\mu := \min\{\hat{\mu}, 1\}$ . Then, there exists  $N$  such that for all  $n > N$   $l_n := \left(\frac{\mu}{2}, \alpha_n\right)$  is accepted. Lemma 2 of Cabrales *et al.* (2014) proves that as  $\frac{1}{n}$  approaches 0, so does the scale of the optimal investment  $\|b^n\|$ . Therefore, for  $\frac{1}{n}$  small enough,  $w - \frac{\mu}{2} + b_k^n$  is in a small environment of  $w - \frac{\mu}{2} < w$  for all  $k$ , a contradiction.

For the second step, from the discussion above it follows that for  $\frac{1}{n}$  small enough,  $w - \mu_n + b_k^n$  is in a  $\delta$ -environment of  $w$  for all  $k$ , if  $a = (\mu, \alpha) \in \mathcal{A}_\epsilon$  is accepted.  $\rho_u(w)$  is continuous, and so for every  $\gamma > 0$  there exists a  $\delta > 0$  small enough such that  $x \in (w - \delta, w + \delta)$  implies  $|\rho_u(x) - \rho_u(w)| < \gamma$ .

For the final step, choose a small positive number  $\eta$ , and consider the CARA agents with absolute risk aversion coefficients  $\rho_u(w) + \eta$  and  $\rho_u(w) - \eta > 0$ . For a small enough environment of  $w$ ,  $I$ ,

$$\rho_u(w) - \eta \leq \inf_{x \in I} \rho_u(x) \leq \sup_{x \in I} \rho_u(x) \leq \rho_u(w) + \eta.$$

This, in turn, implies, using Theorem 3 of Cabrales *et al.* (2014) and a slightly modified version of their Theorem 2, that the coefficient of local taste for  $A$ -informativeness of  $u$  with wealth  $w$  is equal to  $\rho_u^{-1}(w)$ , and that  $R_A(u, w) = S_A(u, w)$ .

(ii) Cabrales *et al.* (2013) showed that  $a = (\mu, \alpha)$  is accepted by an agent

with log utility function if and only if  $I_e(\alpha) > \log\left(\frac{w}{w-\mu}\right)$ . Using a Taylor approximation yields

$$\log\left(\frac{w}{w-\mu}\right) = \log(w) - \log(w-\mu) \approx \frac{1}{w}\mu + \frac{\mu^2}{2w^2}.$$

As shown above, if  $a_n = (\mu_n, \alpha_n) \in \mathcal{A}_{\frac{1}{n}}$  are accepted it must be the case that  $\mu_n \xrightarrow{n \rightarrow \infty} 0$ . It is therefore the case that for  $n$  large enough (when posteriors are close to the prior),  $a_n$  is accepted by agents with log utility function if

$$J_e(a_n) = \frac{I_e(\alpha_n)}{\mu_n} > \frac{1}{w} + O(\mu_n) \xrightarrow{n \rightarrow \infty} \frac{1}{w} = \rho_{\log}(w),$$

and rejected if

$$J_e(a_n) = \frac{I_e(\alpha_n)}{\mu_n} < \frac{1}{w} + O(\mu_n) \xrightarrow{n \rightarrow \infty} \frac{1}{w} = \rho_{\log}(w).$$

For any  $x \in \mathbb{R}_+$ ,  $\frac{1}{x} \equiv w \in \mathbb{R}_+$  satisfies  $\rho_{\log}(w) = x$ , and so by properly translating the log utility function (and changing all but an environment of the baseline wealth level of the agent), one can use a “sandwich” argument of the form used above to complete the proof.  $\square$

### A.1.17 Theorem 1.21

*Proof.* The proof uses the same techniques used above. If  $\rho_u(w) > \rho_v(w')$  then there exists some  $\gamma > 0$  such that  $\rho_u(w) > (1 + \gamma) \cdot \rho_v(w')$ . Following the arguments used before, for  $\epsilon > 0$  small enough, if  $u$  accepts  $a = (\mu, \alpha) \in \mathcal{A}_\epsilon$  then  $v$  accepts  $((1 + \frac{\gamma}{2}) \cdot \mu, \alpha)$ . Together with local consistency and homogeneity this implies that the coefficient of local taste for  $Q$ -informativeness of  $u$  at  $w$  is smaller than the coefficient of local taste for  $Q$ -informativeness of  $v$  at  $w'$ ,

and that  $v$  at  $w$  has at least as much taste for  $Q$ -informativeness as  $u$  at  $w'$ .

4

In the other direction, assume  $\rho_u(w) = \rho_v(w')$ , and by way of contradiction assume that the coefficient of local taste for  $Q$ -informativeness of  $u$  at  $w$  is not equal to the coefficient of local taste for  $Q$ -informativeness of  $v$  at  $w'$ . Without loss of generality, assume that the coefficient of local taste for  $Q$ -informativeness of  $u$  at  $w$  is greater than the coefficient of local taste for  $Q$ -informativeness of  $v$  at  $w'$ . This means that there exists a sequence  $\{a_n\}_{n=1}^{\infty}$  of information transactions, such that for every  $n$ ,  $a_n = (\mu_n, \alpha_n)$  satisfies (a)  $a_n \in \mathcal{A}_{\frac{1}{n}}$ , (b) For some small  $\gamma > 0$ ,  $((1+\gamma) \cdot \mu_n, \alpha_n)$  is accepted by  $u$  at  $w$ , and (c)  $a_n$  is rejected by  $v$  at  $w'$ . But this implies that  $A$  violates local consistency, a contradiction, and so the coefficient of local taste for  $Q$ -informativeness of  $u$  at  $w$  is equal to the coefficient of local taste for  $Q$ -informativeness of  $v$  at  $w'$ . This, in turn, implies that  $u$  at  $w$  has at least as much taste for  $Q$ -informativeness as  $v$  at  $w'$  (and vice versa).  $\square$

### A.1.18 Theorem 1.22

*Proof.* For (i), let  $\delta := \frac{1}{2} \min_i \{\min \{p_i, 1 - p_i\}\}$ . Define

$$B(a) = \begin{cases} A(a) & \|p - q^s\| < \delta \ \forall s \\ \frac{1}{\mu^k} \cdot f(\alpha) & else \end{cases}$$

for some positive  $f$ . Then  $B$  satisfies the required properties since for local transactions (ones with posteriors close to the prior) it is equal to  $A$ , and

---

<sup>4</sup>For details, see Theorem 1.3.

since both  $A$  and  $\frac{1}{\mu}f(\alpha)$  are homogeneous and changes in the price do not change the distance of the posteriors from the prior (and hence the rule that governs  $B$ ). Choosing  $f \equiv 1$  (or many other choices) completes the proof of (ii).  $\square$

### A.1.19 Theorem 1.24

*Proof.* First, observe that for any CARA utility function  $u$  it must be the case that  $Q$ -informativeness is globally at least as attractive for  $u$  as it is for  $u$ , by reflexivity and the lack of wealth effects in CARA functions. Now consider two information transactions,  $a$  and  $a'$ , with  $A(a) > A(a')$ . Consider  $u$  CARA with  $\rho_u \equiv \frac{A(a)+A(a')}{2}$ .  $u$  accepts  $a$  and rejects  $a'$ , implying that  $Q(a) \geq Q(a')$ , since otherwise strong global consistency will be violated (the violation would be the fact that  $Q$ -informativeness is globally at least as attractive for  $u$  as is for itself,  $u$  accepts  $a$  with  $Q(a') > Q(a)$ , but rejects  $a'$ ).

Next, I claim that if  $A(a) > A(a')$ , but  $Q(a) = Q(a')$ , then there exists a transaction  $a_\epsilon$  such that  $A(a_\epsilon) > A(a')$ , but  $Q(a_\epsilon) < Q(a')$  in contradiction to the above result. To see this denote  $a_\epsilon := (\mu + \epsilon, \alpha)$ , where  $a = (\mu, \alpha)$ , and note that from monotonicity of  $Q$ , for any small  $\epsilon > 0$ ,  $Q(a_\epsilon) < Q(a)$ , and from continuity of  $A$ , for small enough  $\epsilon$ ,  $A(a_\epsilon) > A(a')$ .

Finally, I claim that if  $A(a) = A(a')$ , but  $Q(a) > Q(a')$ , then there exists a transaction  $a_\epsilon$  such that  $A(a_\epsilon) < A(a')$ , but  $Q(a_\epsilon) > Q(a')$ . To see this, apply the same argument from the previous paragraph, only this time use the continuity of  $Q$  and the monotonicity of  $A$ .

The upshot of the above discussion is that  $A(a) > A(a') \iff Q(a) >$

$Q(a')$  as required.

□

## Appendix B

# Appendix to Chapter 3

### B.1 Proofs

**Proof of Remark 3.1.** Observe first that any matching that is stable with respect to  $(P_w^k, P_{-w})$  and matches woman  $w$  to a man must be stable with respect to  $P$ , and that any matching  $\tilde{\mu}$  that is stable with respect to  $P$  with  $\tilde{\mu}(w)$  ranked  $\leq k$  must be stable with respect to  $(P_w^k, P_{-w})$ . Hence, setting

$$\mathcal{M}_1 = \{m \in \mathcal{M} \mid m \text{ achievable for } w \text{ under } (P_w^k, P_{-w})\}$$

and

$$\mathcal{M}_2 = \{m \in \mathcal{M} \mid m \text{ achievable for } w \text{ under } P \text{ and } m \text{ ranked } \leq k \text{ in } w\text{'s list}\}$$

we have  $\mathcal{M}_1 = \mathcal{M}_2$ . By the Gale-Shapley result,  $\mu^M[P_w^k, P_{-w}](w)$  is  $w$ 's least preferred element of  $\mathcal{M}_1$ , and hence of  $\mathcal{M}_2$ . Should both sets be empty, then  $\mu^M[P_w^k, P_{-w}](w) = w$ .  $\square$

**Proof of Proposition 3.1.** We introduce Algorithm 1 below and prove that given the same input, Algorithm 1 and the Divorcing Algorithm generate the same output, which in each case is the *MP-DA* outcome described in the statement of the proposition.

Like the Divorcing Algorithm, Algorithm 1 takes as its input a profile  $P$  of preference lists, a woman  $w$ , and a truncation point  $k \in \{1, \dots, N\}$ , and outputs a matching. Algorithm 1 is adapted from an algorithm due to McVitie and Wilson, which differs from *MP-DA* in that the men make offers one at a time instead of in rounds, but is nevertheless outcome equivalent (McVitie and Wilson, 1970). Algorithm 1 is identical to McVitie and Wilson's, except that we explicitly delay selecting man  $\mu^M[P]$  until absolutely necessary. By McVitie and Wilson (1970), the algorithm plainly produces  $\mu^M[P_w^k, P_{-w}]$ , the *MP-DA* outcome when  $w$   $k$ -truncates her preference list.

### Algorithm 1

- **Step 0.** Initialization. Identify the least preferred achievable mate for woman  $w$  under  $(P_w, P_{-w})$  and call this man  $m_l$ . For example, we may identify this man by running *MP-DA*, setting  $m_l = \mu^M[P_w, P_{-w}](w)$ .

Iteration over steps 1 and 2. Preferences in these steps are given by  $(P_w^k, P_{-w})$ .

- **Step 1.** Pick any single man other than  $m_l$  who has not exhausted his preference list. If no such man exists, pick  $m_l$ . If we have picked  $m_l$ , and  $m_l$  is not single, or if  $m_l$  has exhausted his preference list, terminate.

- **Step 2.** The man chosen in the previous step makes an offer to the most preferred woman on his preference list who has not already rejected him. If this woman finds the man acceptable and prefers him to her current mate (or if she is single), she holds his offer and divorces her previous mate (if necessary). Return to step 1.

Let  $\mu^1(P, k, w)$  be the output of Algorithm 1 and recall that  $\mu^{DIV}(P, k, w)$  is the output of the divorcing algorithm.

To establish outcome equivalence of the algorithms, begin by letting  $l$  be the rank of  $w$ 's least preferred achievable mate  $m_l$  under  $P$ .

- If  $k \geq l$ , both algorithms clearly produce  $\mu^M$ , the men-optimal matching under  $P$ .
- If  $k < l$ , then the algorithms will reach a point where they coincide. That is, there will be a point where the sequences of single men chosen coincide, as do the temporary matchings and preference lists.

In Algorithm 1, we claim that (1) at some point,  $m_l$  will make an offer to  $w$ , which will be rejected. (2) From this point forward, the algorithm coincides with the Divorcing Algorithm, just after its initialization step.

1. Under *MP-DA*, when  $w$   $k$ -truncates her list, men are (weakly) worse off than if she reports truthfully (see Gale and Sotomayor (1985)). This means that in  $\mu^M[P_w^k, P_{-w}]$ ,  $m_l$  must be matched with a candidate worse than  $w$ , or possibly with no woman at all. Hence, in Algorithm 1,  $m_l$  must have made an offer to  $w$  (since



he makes offers from his list in order of preference), and this offer must have been rejected.

2. When in Algorithm 1,  $m_l$  makes his offer to  $w$ , no better ranked man has yet done so. Otherwise, let  $m'$  be the first man ranked higher than  $m_l$  to make an offer to  $w$  and backtrack to the point in the algorithm where this offer is made. Note that up to this point, the path of the algorithm is consistent with  $w$  having  $l$ -truncated her preferences, since she has not faced any man ranked  $k$  through  $l$ . But this implies that if  $w$   $l$ -truncated her list, she would receive a mate at least as good as  $m'$ , not  $m_l$ . This contradicts Proposition 3.1.

By the choice-of-proposer rule in the algorithm, we know that when  $m_l$  proposes to  $w$ , he must be the only single man who has not yet exhausted his list. If  $w$  accepted  $m_l$ 's offer, the path of the algorithm would be consistent with  $w$  having  $l$ -truncated her list, and the algorithm would terminate with matching  $\mu^M$ . Hence, by instead rejecting  $m_l$ , we arrive at exactly the position of the Divorcing Algorithm, following step 0.

Thereafter, the algorithms coincide, thus yielding identical outcomes.  $\square$

**Proof of the Lemma in Section 3.3.3.** For each  $i$  and  $k$ , define

$$\mathcal{P}_2^i(k) \equiv \{P_{-w} \mid P_{-w} \in \mathcal{P}_2(k), \mu^M[P_w^k, P_{-w}](w) = m_i\}.$$

We wish to show that  $w$  finds  $\mathcal{P}_2^i(k)$  and  $\mathcal{P}_2^j(k)$  equally probable, for all  $k$  and all  $i, j \leq k$ . We proceed by finding a bijection from  $\mathcal{P}_2^i(k)$  to  $\mathcal{P}_2^j(k)$

which is probability preserving with respect to  $w$ 's beliefs.

For  $i, j \in \{1, \dots, k\}$ , we define a mapping  $f_{ijk} : \mathcal{P}_2(k) \rightarrow \mathcal{P}_2(k)$ . Let  $f_{ijk}(P_{-w}) \equiv P'_{-w}$  be given by the following:

1. Switch  $m_i$  and  $m_j$  everywhere. Switch the positions of  $m_i$  and  $m_j$  in each woman's list, and swap  $m_i$  and  $m_j$ 's preference lists (this is like relabeling).
2. Switch back  $m_i$  and  $m_j$  in  $w$ 's list.

Notice that this is equivalent to swapping  $m_i$  and  $m_j$  in  $w$ 's list only, and then relabeling  $i$  and  $j$ .<sup>1</sup>

Suppose  $P_{-w} \in \mathcal{P}_2^i(k)$ . The fact that  $w$  finds  $P_{-w}$  and  $P'_{-w}$  equally probable follows directly from the definition of  $\mathcal{M}$ -symmetry. We will show that  $P'_{-w} \in \mathcal{P}_2^j(k)$ . Note that it is not immediately clear that we even have  $P'_{-w} \in \mathcal{P}_2(k)$ , that is, that under  $P'_{-w}$ ,  $k$ -truncation still yields an improvement for  $w$ .

We think of the matching as arising from *MP-DA*. Since  $P_{-w} \in \mathcal{P}_2(k)$ , if  $w$  does not truncate, she will be matched with a man worse than  $m_k$ . Hence, during the process of the algorithm, she will not receive an offer from any man  $m_1, \dots, m_k$ . Hence, rearranging these men in  $w$ 's list will not affect the outcome, and in particular, swapping  $m_i$  and  $m_j$  will not affect the outcome (the stable matching). Furthermore, since  $P_{-w} \in \mathcal{P}_2^i(k)$  we know that under  $P_{-w}$ ,  $k$ -truncation leaves  $w$  matched with  $m_i$ . Using proposition 3.1, we know that during the chain of proposals following an “ex-post”  $k$ -truncation by  $w$ ,

---

<sup>1</sup>For the trivial case,  $i = j$ , we use the identity mapping.

the first man to make an offer to  $w$  will be  $m_i$ . Hence, this will still be true if  $w$  swaps the position of  $m_i$  and  $m_j$  in her list.

Thus, we have that if  $w$  switches  $m_i$  and  $m_j$  in her list,  $k$ -truncation will yield an improvement and she will again be matched with  $m_i$ . But now relabeling  $m_i$  and  $m_j$  (so that  $w$ 's list is  $(m_1, m_2, m_3, \dots)$ ), we have that  $P'_{-w} \in \mathcal{P}_2^j(k)$ .

Hence,  $f_{ijk}(\cdot)$  is a bijection from  $\mathcal{P}_2^i(k)$  to  $\mathcal{P}_2^j(k)$ , which is probability preserving with respect to  $w$ 's beliefs. This is sufficient to prove the proposition.  $\square$

To prove Theorem 4.1, we begin with a lemma demonstrating that even upon submitting a vanishingly small truncation of one's list (relative to the length of one's full preference list), we still see gains relative to truthful reporting. We examine the case where a woman's payoff is given by her partner rank, and being unmatched is treated as rank  $N + 1$ . At the end of the proof, we show that the result also holds for the more general preferences described in the statement of the Theorem 4.1.<sup>2</sup>

**Lemma B.1.** *There exists  $N^*$  such that for every  $N > N^*$ , the gain to woman  $w$  from truncating at  $7 \log^2 N$  relative to truthful reporting is strictly greater than zero. Furthermore, the expected rank of  $w$ 's mate is lower than (better than) her expected mate rank from truthful reporting by at least  $\frac{1+N}{2+\log N} - 7 \log^2 N - 2$ .*

---

<sup>2</sup>Throughout the proofs of Lemma B.1, Lemma B.2, and Theorem 4.1, for any fractional  $x \in \mathbb{R}^+$ , we treat  $x$ -truncation as  $\lfloor x \rfloor$ -truncation.

**Proof of Lemma 1.** First, in the case of no truncation, we know from Pittel (1989, p. 545) that the expected rank of  $w$ 's husband,  $R_w(\mu_m)$  satisfies:

$$\mathbb{E}[R_w] \geq \frac{1+N}{1+H_N} \geq \frac{1+N}{2+\log N},$$

where  $H_N = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{N}$ , the  $N$ th harmonic number. Let  $D$  be the highest (worst) rank some woman gets under *WP-DA* when all agents report their preferences truthfully. Using Theorem 6.1 from Pittel (1992), we observe that for  $N$  large enough,

$$\Pr(D \leq 7 \log^2 N) \geq 1 - \frac{1}{N}.$$

Therefore, truncating at  $7 \log^2 N$  ensures an expected rank of at most  $7 \log^2 N \times (1 - \frac{1}{N}) + \frac{1}{N} \times (N+1)$ . Hence, the expected gain (in rank) from truncation,  $\Delta$ , satisfies:

$$\Delta \geq \frac{1+N}{2+\log N} - 7 \log^2 N - 2.$$

The right hand side approaches infinity as  $N$  grows to infinity, so for  $N$  large enough,  $\Delta > 0$ . □

In Lemma B.1 we have established that truncating at  $7 \log^2 N$  ensures a gain (in terms of expected partner rank) relative to truthful revelation that grows arbitrarily large as  $N \rightarrow \infty$ . Note that this gain is an absolute measure. As measured as a *fraction* of the expected payoff from truthful revelation, the gains from truncation go to zero.<sup>3</sup>

---

<sup>3</sup>Recall that as a fraction of  $N$ , the expected partner rank for women (as well as for men) converges to 0 as  $N$  grows large.

It remains to establish that as a fraction of the market size  $N$ , the degree of *optimal* truncation goes to 0. To do this, we will first show that any truncation of a constant fraction of one's list is (asymptotically) outperformed by the level of truncation found in Lemma B.1.

**Lemma B.2.** *For any fraction  $\alpha \in (0, 1)$ ,*

- i) there exists  $N(\alpha)$  such that for every  $N > N(\alpha)$ , a woman's payoff from truncating at  $\alpha N$  is lower than that from truncating at  $7 \log^2 N$ ;*
- ii) there exists  $N^*(\alpha)$  such that for every  $N > N^*(\alpha)$ , a woman's payoff from truncating at  $xN$  is lower than that from truncating at  $7 \log^2 N$  for every  $x \in [\alpha, 1]$ .*

**Proof of Lemma 2.** We begin by proving *i)* for the case of  $\alpha = 1 - \frac{1}{e}$ .

Let  $\bar{\Delta}$  be the expected difference between the rank of the mate under truncation at  $\alpha N$  and the rank of the mate under truthful revelation. Let  $\epsilon > 0$  be a small number. Let  $A_N$  be the event  $\{w \text{ gets fewer than } (1 - \epsilon) \log N \text{ offers, or else more than } (1 + \epsilon) \log N, \text{ before } MP\text{-}DA \text{ stops}\}$ . Let  $P_N = P(A_N)$ . We then have:

$$\begin{aligned}
\bar{\Delta} &\leq P_N \times N + (1 - P_N) \times \Pr\{\text{Rank}(\mu^M(w)) > (1 - \frac{1}{e})N \mid \neg A_N\} \times N \\
&\leq P_N \times N + \Pr\{\text{Rank}(\mu^M(w)) > (1 - \frac{1}{e})N \mid \neg A_N\} \times N \\
&\leq P_N \times N + \left(\frac{1}{e}\right)^{(1-\epsilon) \log N} \times N.
\end{aligned}
\tag{B.1.0.1}$$

Note that truncation may only matter in the event  $\{\text{Rank}(\mu_M(w)) > (1 - \frac{1}{e})N\}$ , which is included in the event  $B = \{\{\text{Rank}(\mu_M(w)) > (1 - \frac{1}{e})N\} \cap \neg A_N\} \cup A_N$ . In the first inequality, we have replaced the conditional benefits from

truncation, be they positive or negative, with  $N$ , and considered the event  $B$ . In the last inequality, we treat offers  $w$  receives as independent draws (and invoke the Principle of Deferred Decisions), when in fact the draws are “without replacement,” which would yield a lower probability. These substitutions are all acceptable as we are finding an upper bound on  $\bar{\Delta}$ .

Using Equation 2.16 from Pittel *et al.* (2007), we know that there exists some  $c > 0$  such that for  $N$  large enough,  $P_N \leq \exp\left(-c \cdot \log^{\frac{1}{3}} N\right)$ . Hence,

$$\begin{aligned} \bar{\Delta} &\leq \exp\left(-c \cdot \log^{\frac{1}{3}} N\right) \times N + \left(\frac{1}{e}\right)^{(1-\epsilon) \log N} \times N \\ &= \frac{N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} + N^\epsilon. \end{aligned} \tag{B.1.0.2}$$

We now must show that for large  $N$ ,  $\frac{N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} + N^\epsilon \leq \frac{1+N}{2+\log N} - 7 \log^2 N - 2$ , which by Lemma B.1 will imply  $\Delta \geq \bar{\Delta}$ .

We have  $\frac{N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} + N^\epsilon \leq \frac{2N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)}$  for  $N$  large enough, since  $N^\epsilon \leq \frac{N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} \iff 1 \leq \frac{N^{1-\epsilon}}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} \iff c \cdot \log^{\frac{1}{3}} N \leq (1-\epsilon) \log N$ , which clearly holds for large  $N$ .

Hence, it is sufficient to prove that:

$$\frac{2N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} \leq \frac{1+N}{2+\log N} - 7 \log^2 N - 2.$$

Since for large  $N$ ,  $7 \log^2 N + 2 < N^\epsilon$ , it suffices to show that  $\frac{3N}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} \leq \frac{1+N}{2+\log N}$ , which is implied by  $\frac{3N+3}{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)} \leq \frac{1+N}{2+\log N} \iff 3 \leq \frac{\exp\left(c \cdot \log^{\frac{1}{3}} N\right)}{2+\log N}$ .

Observe that

$$\lim_{N \rightarrow \infty} \frac{\exp\left(c \log^{\frac{1}{3}} N\right)}{2 + \log N} = \lim_{x \rightarrow \infty} \frac{\exp(cx)}{2 + x^3} = \infty$$

since  $c$  is greater than 0. This completes the proof for the case of  $(1 - \frac{1}{e})N$ .

To show that *i*) holds, we now consider general  $\alpha \in (0, 1)$ . Let  $r \equiv \frac{1}{1-\alpha} > 1$ , so that  $1 - \frac{1}{r} = \alpha$ . An analogous proof holds with truncation at  $(1 - \frac{1}{r})N$ . Probability  $P_N$  will remain unchanged, and in Equation B.1.0.1, instead of  $(\frac{1}{e})^{(1-\epsilon) \log N}$  we have  $(\frac{1}{r})^{(1-\epsilon) \log N} = (\frac{1}{e})^{\log r \cdot (1-\epsilon) \log N} = (\frac{1}{N})^{(1-\epsilon) \log r} = (\frac{1}{N})^{\delta(\alpha)}$ , where  $\delta(\alpha) \equiv (1 - \epsilon) \log r = (1 - \epsilon) \log \frac{1}{1-\alpha} > 0$ . We may then replace  $N^\epsilon$  with  $N^{1-\delta(\alpha)}$  in Equation B.1.0.2, and the remaining argument will hold.

To show that *ii*) holds, observe that the critical appearance of  $\alpha$  is in the inequality  $N^{1-\delta(\alpha)} \leq \frac{N}{\exp\left(c \log^{\frac{1}{3}} N\right)}$ . For every  $x > \alpha$ , we have  $N^{1-\delta(x)} \leq N^{1-\delta(\alpha)}$ . Hence, for any  $N$  large enough so that  $N^{1-\delta(\alpha)} \leq \frac{N}{\exp\left(c \log^{\frac{1}{3}} N\right)}$ , we have that  $N^{1-\delta(x)} \leq \frac{N}{\exp\left(c \log^{\frac{1}{3}} N\right)}$  holds as well, demonstrating *ii*).  $\square$

**Proof of Theorem 4.1.** By way of contradiction, assume that  $\lim_{N \rightarrow \infty} \frac{k^*(N)}{N} = 0$  does not hold. This implies that there exists a subsequence  $\{N_j\}$  such that  $\lim_{j \rightarrow \infty} \frac{k^*(N_j)}{N_j} = b > 0$ , so for  $N_j$  large enough,  $\frac{k^*(N_j)}{N_j} > b/2$ . By Lemma B.2 *ii*), we know that for large enough  $N_j$ , truncating at  $7 \log^2 N_j$  outperforms truncating at  $xN_j$  for any  $x \geq b/2$ . But this contradicts the optimality of the truncations at  $k^*(N_j)$ , and so concludes the proof for the case when payoffs are given by partner rank. By applying Corollary 4.2, we see that the result also holds for any strictly increasing, convex transformation of such preferences.  $\square$

To prove Theorems 3.2 and 3.4, we show that the following lemma holds:

**Lemma B.3.** *Let  $\tau$  be a profile of strategies where each man reports truthfully and women play truncation strategies. Let  $\sigma^*$  be an equilibrium in truncation strategies, such that every woman in  $\mathcal{W} \setminus \{w\}$  truncates more at any state of the world (in the sense of FOSD) and men report truthfully. Then woman  $w$  is weakly better off under  $\sigma^*$  than under  $\tau$ .*

**Proof of Lemma B.3.** Since other women truncate more under  $\sigma^*$ , it is clear that the payoff to  $w$  from the profile  $(\tau_w, \sigma_{-w}^*)$  is weakly higher than her payoff under  $\tau$ . Moreover, since  $\sigma_w^*$  is a best response to  $\sigma_{-w}^*$ , the payoff to  $w$  from  $(\sigma_w^*, \sigma_{-w}^*)$  is weakly greater than that under  $(\tau_w, \sigma_{-w}^*)$ .  $\square$

**Proof of Theorem 3.2.** *i)* is a direct consequence of Lemma B.3. Proofs for *ii)* and *iii)* were given in-text.

**Proof of Theorem 4.** *i)* follows from symmetry. *ii)* is a direct consequence of Lemma B.3.

**Proof of Proposition 4.1.** To prove the proposition, we first show that if under low correlation, we prefer truncating less to more, than under high correlation we definitely prefer truncating less to more.



First, observe that

$$\begin{aligned}
\mathbb{E}_{p^{\alpha'}}[v(k, \tilde{P}_{-w})] &= \sum_{P_{-w}} p^{\alpha'}(P_{-w}) v(k, P_{-w}) \\
&= \sum_{P_{-w}} [(1 - \alpha')p(P_{-w}) + \alpha'p^C(P_{-w})] v(k, P_{-w}) \\
&= \sum_{P_{-w}} \left[ (1 - \alpha')p(P_{-w}) + \alpha \frac{1 - \alpha'}{1 - \alpha} p^C(P_{-w}) + \frac{\alpha' - \alpha}{1 - \alpha} p^C(P_{-w}) \right] v(k, P_{-w}) \\
&= \left( \frac{1 - \alpha'}{1 - \alpha} \right) \mathbb{E}_{p^{\alpha}}[v(k, \tilde{P}_{-w})] + \left( \frac{\alpha' - \alpha}{1 - \alpha} \right) \mathbb{E}_{p^C}[v(k, \tilde{P}_{-w})].
\end{aligned}$$

Now suppose that for  $k_1, k_2 \in \{1, \dots, N\}$  with  $k_2 > k_1$ , we have

$$\mathbb{E}_{p^{\alpha}}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^{\alpha}}[v(k_1, \tilde{P}_{-w})]. \quad (\text{B.1.0.3})$$

Then since

$$\mathbb{E}_{p^C}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^C}[v(k_1, \tilde{P}_{-w})],$$

we must have

$$\mathbb{E}_{p^{\alpha'}}[v(k_2, \tilde{P}_{-w})] \geq \mathbb{E}_{p^{\alpha'}}[v(k_1, \tilde{P}_{-w})]. \quad (\text{B.1.0.4})$$

If the inequality in (B.1.0.3) is strict, then so too is the inequality in (B.1.0.4).

We can now use this payoff comparative static to sort optimal truncation points as follows.

By definition,  $k_l(\alpha, p, u_w)$  satisfies

$$\mathbb{E}_{p^{\alpha}}[v(k_l(\alpha, p, u_w), \tilde{P}_{-w})] > \mathbb{E}_{p^{\alpha}}[v(k, \tilde{P}_{-w})] \quad \forall k < k_l(\alpha, p, u_w).$$

From (B.1.0.4), we must then have

$$\mathbb{E}_{p^{\alpha'}}[v(k_l(\alpha, p, u_w), \tilde{P}_{-w})] > \mathbb{E}_{p^{\alpha'}}[v(k, \tilde{P}_{-w})] \quad \forall k < k_l(\alpha, p, u_w),$$

so that  $k_l(\alpha', p, u_w) \geq k_l(\alpha, p, u_w)$ .

Similarly,  $k_h(\alpha, p, u_w)$  satisfies

$$\mathbb{E}_{p^\alpha}[v(k_h(\alpha, p, u_w), \tilde{P}_{-w})] \geq \mathbb{E}_{p^\alpha}[v(k, \tilde{P}_{-w})] \quad \forall k < k_h(\alpha, p, u_w).$$

From (B.1.0.4), we must then have

$$\mathbb{E}_{p^{\alpha'}}[v(k_h(\alpha, p, u_w), \tilde{P}_{-w})] \geq \mathbb{E}_{p^{\alpha'}}[v(k, \tilde{P}_{-w})] \quad \forall k < k_h(\alpha, p, u_w),$$

so that  $k_h(\alpha', p, u_w) \geq k_h(\alpha, p, u_w)$ . □